Morphometric, behavioral, and genomic evidence for a new Orangutan species

Nater, Alexander ; Mattle-Greminger, Maja Patricia ; Nurcahyo, Anton ; Nowak, Matthew G ; de Manuel, Marc ; Desai, Tariq ; Groves, Colin ; Pybus, Marc ; Sonay, Tugce Bilgin ; Roos, Christian ; Lameira, Adriano R ; Wich, Serge A ; Askew, James ; Davila-Ross, Marina ; Fredriksson, Gabriella M ; de Valles, Guillem ; Casals, Ferran ; Prado-Martinez, Javier ; Goossens, Benoit ; Verschoor, Ernst J ; Warren, Kristin S ; Singleton, Ian ; Marques, David A ; Pamungkas, Joko ; Perwitasari-Farajallah, Dyah ; Rianti, Puji ; Tuuga, Augustine ; Gut, Ivo G ; Gut, Marta ; Orozco-terWengel, Pablo ; van Schaik, Carel P ; Bertranpetit, Jaume ; Anisimova, Maria ; Scally, Aylwyn ; Marques-Bonet, Tomas ; Meijaard, Erik ; Krützen, Michael

Abstract: Six extant species of non-human great apes are currently recognized: Sumatran and Bornean orangutans, eastern and western gorillas, and chimpanzees and bonobos [1]. However, large gaps remain in our knowledge of fine-scale variation in hominoid morphology, behavior, and genetics, and aspects of great ape taxonomy remain in flux. This is particularly true for orangutans (genus: Pongo), the only Asian great apes and phylogenetically our most distant relatives among extant hominids [1]. Designation of Bornean and Sumatran orangutans, P. pygmaeus (Linnaeus 1760) and P. abelii (Lesson 1827), as distinct species occurred in 2001 [1, 2]. Here, we show that an isolated population from Batang Toru, at the southernmost range limit of extant Sumatran orangutans south of Lake Toba, is distinct from other northern Sumatran and Bornean populations. By comparing cranio-mandibular and dental characters of an orangutan killed in a human-animal conflict to those of 33 adult male orangutans of a similar developmental stage, we found consistent differences between the Batang Toru individual and other extant Ponginae. Our analyses of 37 orangutan genomes provided a second line of evidence. Model-based approaches revealed that the deepest split in the evolutionary history of extant orangutans occurred 3.38 mya between the Batang Toru population and those to the north of Lake Toba, whereas both currently recognized species separated much later, about 674 kya. Our combined analyses support a new classification of orangutans into three extant species. The new species, Pongo tapanuliensis, encompasses the Batang Toru population, of which fewer than 800 individuals survive. VIDEO ABSTRACT.

DOI: https://doi.org/10.1016/j.cub.2017.09.047

Casals, Ferran; Prado-Martinez, Javier; Goossens, Benoit; Verschoor, Ernst J; Warren, Kristin S; Singleton, Ian; Marques, David A; Pamungkas, Joko; Perwitasari-Farajallah, Dyah; Rianti, Puji; Tuuga, Augustine; Gut, Ivo G; Gut, Marta; Orozco-terWengel, Pablo; van Schaik, Carel P; Bertranpetit, Jaume; Anisimova, Maria; Scally, Aylwyn; Marques-Bonet, Tomas; Meijaard, Erik; Krützen, Michael (2017). Morphometric, behavioral, and genomic evidence for a new Orangutan species. Current Biology, 27(22):3487-3498.e10.
DOI: https://doi.org/10.1016/j.cub.2017.09.047

Report

Title: Morphometric, behavioral, and genomic evidence for a new orangutan species

Authors: Alexander Nater[1,2,3§]*, Maja P. Mattle-Greminger[1,2§], Anton Nurcahyo[4§], Matthew G. Nowak[5,6§], Marc de Manuel[7], Tariq Desai[8], Colin Groves[4], Marc Pybus[7], Tugce Bilgin Sonay[1], Christian Roos[9], Adriano R. Lameira[10,11], Serge A. Wich[12,13], James Askew[14], Marina Davila-Ross[15], Gabriella Fredriksson[5,13], Guillem de Valles[7], Ferran Casals[16], Javier Prado-Martinez[17], Benoit Goossens[18,19,20,21], Ernst J. Verschoor[22], Kristin S. Warren[23], Ian Singleton[5,24], David A. Marques[1,25], Joko Pamungkas[26,27], Dyah Perwitasari-Farajallah[26,28], Puji Rianti[28,26,1], Augustine Tuuga[20], Ivo G. Gut[29,30], Marta Gut[29,30], Pablo Orozco-terWengel[18], Carel P. van Schaik[1], Jaume Bertranpetit[7,31], Maria Anisimova[32,33], Aylwyn Scally[8], Tomas Marques-Bonet[7,29,34], Erik Meijaard[4,35]* and Michael Krützen[1]*

[§]These authors contributed equally to this work.

*Correspondence to: michael.kruetzen@aim.uzh.ch (MK, lead contact), alexander.nater@uzh.ch (AlN), emeijaard@gmail.com (EM),

Affiliations:

[1]Evolutionary Genetics Group, Department of Anthropology, University of Zurich, Winterthurerstrasse 190, 8057 Zürich, Switzerland.

[2]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057 Zürich, Switzerland.

[3]Lehrstuhl für Zoologie und Evolutionsbiologie, Department of Biology, University of Konstanz, Universitätsstrasse 10, 78457 Konstanz, Germany.

[4]School of Archaeology and Anthropology, Australian National University, Canberra, Australia.

[5]Sumatran Orangutan Conservation Programme (PanEco-YEL), Jalan Wahid Hasyim 51/74, Medan 20154, Indonesia.

[6]Department of Anthropology, Southern Illinois University, 1000 Faner Drive, Carbondale, IL 62901, USA.

[7]Institut de Biologia Evolutiva (UPF-CSIC), Universitat Pompeu Fabra, Doctor Aiguader 88, Barcelona 08003, Spain.

[8]Department of Genetics, University of Cambridge, Downing Street, Cambridge, CB2 3EH, UK.

[9]Gene Bank of Primates and Primate Genetics Laboratory, German Primate Center, Leibniz Institute for Primate Research, 37077 Göttingen, Germany.

[10]Department of Anthropology, Durham University, Dawson Building, South Road, Durham, DH1 3LE, UK.

[11]School of Psychology & Neuroscience, St. Andrews University, St Mary's Quad, South Street, St. Andrews, Fife, KY16 9JP, Scotland, United Kingdom.

[12]School of Natural Sciences and Psychology, Liverpool John Moores University, James Parsons Building, Byrom Street, L33AF Liverpool, UK.

[13]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Sciencepark 904, Amsterdam 1098, Netherlands.

[14]Department of Biological Sciences, University of Southern California, 3616 Trousdale Parkway, Los Angeles, CA 90089, USA.

[15]Department of Psychology, University of Portsmouth, King Henry Building, King Henry 1st Street, Portsmouth, PO1 2DY, UK.

[16]Servei de Genòmica, Universitat Pompeu Fabra, Doctor Aiguader 88, Barcelona 08003, Spain.

[17]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK.

[18]School of Biosciences, Cardiff University, Sir Martin Evans Building, Museum Avenue, Cardiff CF10 3AX, UK.

[19]Danau Girang Field Centre, c/o Sabah Wildlife Department, Wisma Muis, 88100 Kota Kinabalu, Sabah, Malaysia.

[20]Sabah Wildlife Department, Wisma Muis, 88100 Kota Kinabalu, Sabah, Malaysia.

[21]Sustainable Places Research Institute, Cardiff University, 33 Park Place, Cardiff CF10 3BA, UK.

[22]Department of Virology, Biomedical Primate Research Centre, Lange Kleiweg 161, 2288GJ Rijswijk, The Netherlands.

[23]Conservation Medicine Program, College of Veterinary Medicine, Murdoch University, South Street, Murdoch 6150, Australia.

[24]Foundation for a Sustainable Ecosystem (YEL), Medan, Indonesia.

[25]Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, 3012 Bern, Switzerland.

[26]Primate Research Center, Bogor Agricultural University, Bogor 16151, Indonesia.

[27]Faculty of Veterinary Medicine, Bogor Agricultural University, Darmaga Campus, Bogor 16680, Indonesia.

[28]Animal Biosystematics and Ecology Division, Department of Biology, Bogor Agricultural University, Jalan Agatis, Dramaga Campus, Bogor 16680, Indonesia.

[29]CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Baldiri i Reixac 4, Barcelona 08028, Spain.

[30]Universitat Pompeu Fabra (UPF), Plaça de la Mercè, 10, 08002 Barcelona, Spain.

[31]Leverhulme Centre for Human Evolutionary Studies, Department of Archaeology and Anthropology, University of Cambridge, Cambridge, UK.

[32]Institute of Applied Simulations, School of Life Sciences and Facility Management, Zurich University of Applied Sciences ZHAW, Einsiedlerstrasse 31a, 8820 Wädenswil, Switzerland.

[33]Swiss Institute of Bioinformatics, Quartier Sorge - Batiment Genopode, 1015 Lausanne, Switzerland.

[34]Institucio Catalana de Recerca i Estudis Avançats (ICREA), Barcelona 08010, Spain.

[35]Borneo Futures, Bandar Seri Begawan, Brunei Darussalam.

Summary

Six extant species of non-human great apes are currently recognized: Sumatran and Bornean orangutans, eastern and western gorillas, and chimpanzees and bonobos [1]. However, large gaps remain in our knowledge of the fine-scale variation in hominoid morphology, behavior, and genetics, and aspects of great ape taxonomy remain in flux. This is particularly true for orangutans (genus: *Pongo*), the only Asian great apes, and phylogenetically our most distant relatives among extant hominids [1]. Definitive designation of Bornean and Sumatran orangutans, *P. pygmaeus* (Linnaeus 1760) and *P. abelii* (Lesson 1827), as distinct species took place as recently as 2001 [1, 2]. Based on morphological, behavioral and environmental data, and corroborated by population genetic analyses of 37 orangutan genomes, we show that an isolated population of the Sumatran orangutan south of Lake Toba, Batang Toru, is highly distinct from the northern Sumatran and Bornean populations. The deepest split in the evolutionary history of extant orangutans occurred ~3.38 Ma between this remnant population south of Lake Toba and those to the north, while both currently recognized species separated much later about 674 ka. A morphometric analysis based on cranio-mandibular and dental characters as well as behavioral data revealed consistent differences between individuals from Batang Toru and other extant Ponginae. Our combined analyses support a new classification of orangutans into three extant species. One of them, *P. tapanuliensis*, encompasses the Batang Toru population, of which fewer than 800 individuals survive.

Results and Discussion

Despite decades of field studies [3] our knowledge of variation among orangutans remains limited as many populations occur in isolated and inaccessible habitats, leaving questions regarding their evolutionary history and taxonomic classification largely unresolved. In particular, Sumatran populations south of Lake Toba had long been overlooked, even though a 1939 review of the species' range mentioned that orangutans had been reported in several forest areas in that region [4]. Based on diverse sources of evidence, we describe a new orangutan species, *Pongo tapanuliensis*, which encompasses a geographically and genetically isolated population found in the Batang Toru area at the southernmost range of extant Sumatran orangutans, south of Lake Toba, Indonesia.

Systematics

Genus *Pongo* Lacépède, 1799

Pongo tapanuliensis sp. nov. Nurcahyo, Meijaard, Nowak, Fredriksson & Groves

Tapanuli Orangutan

Etymology. The species name refers to three North Sumatran districts (North, Central, and South Tapanuli) to which *P. tapanuliensis* is endemic.

Holotype. The complete skeleton of an adult male orangutan that died from wounds sustained by local villagers in November 2013 near Sugi Tonga, Marancar, Tapanuli (Batang Toru) Forest Complex (1^{0}35'54.1"N, 99^{0}16'36.5"E), South Tapanuli District, North Sumatra, Indonesia. Skull and postcranium are lodged in the Museum Zoologicum Bogoriense, Indonesia, accession number MZB39182. High-resolution 3D reconstructions of the skull and mandible are available as supplementary material.

Paratypes. Adult individuals of *P. tapanuliensis* (P2591-M435788 – P2591-M435790) photographed by Tim Laman in the Batang Toru Forest Complex (1^{0}41'9.1"N, 98^{0}59'38.1"E), North Tapanuli District, North Sumatra, Indonesia. Paratypes are available from http://www.morphobank.org (Login: 2591 / Password: tapanuliorangutan).

Differential diagnosis. Unless otherwise stated, all units are [mm]. Summary statistics for all measurements are listed in Tables S1–3. *Pongo tapanuliensis* differs from all extant orangutans in the breadth of the upper canine (21.5, vs. <20.86); the shallow face depth (6.0 vs. >8.4); the narrower interpterygoid distance (at posterior end of pterygoids 33.8 vs. >43.9; at anterior end of pterygoids, 33.7 vs. >43.0); the shorter tympanic tube (23.9 vs. >28.4, mostly >30); the shorter temporomandibular joint (22.5 vs. >24.7); the narrower maxillary incisor row (28.3 vs. >30.1); the narrower distance across the palate at the first molars (62.7 vs. >65.7); the shorter horizontal length of

the mandibular symphysis (49.3 vs. >53.7); the smaller inferior transverse torus (horizontal length from anterior surface of symphysis 31.8 compared to >36.0); and the width of the ascending ramus of the mandible (55.9 vs. >56.3).

Pongo tapanuliensis differs specifically from *P. abelii* by its deep suborbital fossa, triangular pyriform aperture, and angled facial profile; the longer nuchal surface (70.5 vs. <64.7); the wider rostrum, posterior to the canines (59.9 vs. <59); the narrower orbits (33.8 vs. <34.6); the shorter (29.2 vs. >30.0) and narrower foramen magnum (23.2 vs. >23.3); the narrower bicondylar breadth (120.0 vs. >127.2); the narrower mandibular incisor row (24.4 vs. >28.3); the greater mesio-distal length of the upper canine (19.44 vs. <17.55). The male long call has a higher maximum frequency range of the roar pulse type (> 800 Hz vs. <747) with a higher 'shape' (>952 Hz/s vs. <934).

Pongo tapanuliensis differs from *P. pygmaeus* by possessing a nearly straight zygomaxillary suture; the lower orbit (orbit height 33.4 vs. >35.3); the male long call has a longer duration (>111 seconds vs. <90) with a greater number of pulses (>52 pulses vs. <45), and is delivered at a greater rate (>0.82 pulses per 20 seconds vs. <0.79).

Pongo tapanuliensis differs specifically from *Pongo 'pygmaeus' palaeosumatrensis* in the smaller size of the first upper molar (mesio-distal length 13.65 vs. >14.0, buccolingual breadth 11.37 vs. >12.10, crown area 155.2 mm^2 vs. >175.45, Figure S1).

Description. Craniometrically, the type skull of *P. tapanuliensis* (Figure 1B) is significantly smaller than any skull of comparable developmental stage of other orangutans; it falls outside of the interquartile ranges of *P. abelii* and *P. pygmaeus* for 24 of 39 cranio-mandibular measurements (Table S1). A PCA of 26 cranio-mandibular measurements commonly used in primate taxonomic classification [5, 6] shows consistent differences between *P. tapanuliensis* and the two currently recognized species (Figs. 1C and S2).

The external morphology of *P. tapanuliensis* is more similar to *P. abelii* in its linear body build and more cinnamon pelage than *P. pygmaeus*. The hair texture of *P. tapanuliensis* is frizzier, contrasting in particular with the long, loose body hair of *P. abelii*. *Pongo tapanuliensis* has a prominent moustache and flat flanges covered in downy hair in dominant males, while flanges of older males resemble more those of Bornean males. Females of *P. tapanuliensis* have beards, unlike *P. pygmaeus*.

Distribution. *Pongo tapanuliensis* occurs only in a small number of forest fragments in the districts of Central, North, and South Tapanuli, Indonesia (Figure 1A). The total distribution covers approximately 1,000 km^2, with an estimated population size of fewer than 800 individuals [7]. The current distribution of *P. tapanuliensis* is almost completely restricted to medium elevation hill and submontane forest (~300–1300 m asl) [7-9]. While densities are highest in primary forest, it does

occur at lower densities in mixed agroforest at the edge of primary forest areas [10, 11]. Until relatively recently, *P. tapanuliensis* was more widespread to the south and west of the current distribution, although evidence for this is largely anecdotal [12, 13].

Other hominoid species and subspecies were previously described using standard univariate and multivariate techniques to quantify morphological character differences (e.g. for genus *Pongo*: [5, 14, 15]; for genus *Pan*: [5, 16, 17]; for genus *Gorilla*: [18, 19]). Here, we used an integrative approach by corroborating the morphological analysis with whole-genome data of 37 orangutans with known provenance, covering the entire range of extant orangutans including areas never sampled before (Figure 2A, Table S4). We applied a model-based approach to statistically evaluate competing demographic models, identify independent evolutionary lineages, and test for the presence or absence of ongoing introgression between sister lineages, enabling us to deal with complex, realistic models of speciation. We did not compare the genetic differentiation among the three species in the genus *Pongo* with that of other hominoids. We deem such an approach problematic and not suitable to evaluate whether *P. tapanuliensis* constitutes a new species, because estimates of genetic differentiation reflect a combination of divergence time, demographic history, and gene flow, and are also potentially influenced by the employed genetic marker system [20-22].

A principal component analysis (PCA; Figure 2B) of genomic diversity highlighted the divergence between individuals from Borneo and Sumatra (PC1), but also separated *P. tapanuliensis* from *P. abelii* (PC2). The same clustering pattern was also found in a model-based analysis of population structure (Figure 2C), and is consistent with an earlier genetic study analyzing a larger number of non-invasively collected samples using microsatellite markers [23]. However, while powerful in detecting extant population structure, population history and speciation cannot be inferred, as they are not suited to distinguish between old divergences with gene flow and cases of recent divergence with isolation [24, 25]. To address this problem and further investigate the timing of population splits and gene flow, we therefore employed different complementary modeling and phylogenetic approaches.

We applied an Approximate Bayesian Computation (ABC) approach, which allows to infer and compare arbitrarily complex demographic modes based on the comparison of the observed genomic data to extensive population genetic simulations [26]. Our analyses revealed three deep evolutionary lineages in extant orangutans (Figs. 3A and B). Colonization scenarios in which the earliest split within *Pongo* occurred between the lineages leading to *P. abelii* and *P. tapanuliensis* were much better supported than scenarios in which the earliest split was between Bornean and Sumatran species (models 1 vs. models 2, combined posterior probability: 99.91%, Figure 3A). Of the two best scenarios, a model postulating colonization of both northern Sumatra and Borneo from an ancestral population likely situated south of Lake Toba on Sumatra, had the highest support (model 1a vs. model 1b, posterior probability 97.56%, Figure 3A). Our results supported a scenario in which

orangutans from mainland Asia first entered Sundaland south of what is now Lake Toba on Sumatra, the most likely entry point based on paleogeographic reconstructions [27]. This ancestral population, of which *P. tapanuliensis* is a direct descendant, then served as a source for the subsequent different colonization events of what is now Borneo, Java and northern Sumatra.

We estimated the split time between populations north and south of Lake Toba at ~3.4 Ma (Figure 3B, Table S5). Under our best-fitting model, we found evidence for post-split gene flow across Lake Toba (~0.3–0.9 migrants per generation, Table S5), which is consistent with highly significant signatures of gene flow between *P. abelii* and *P. tapanuliensis* using D-statistics (CK, BT, WA, *Homo sapiens*: D= -0.2819, p-value<0.00001; WK, BT, LK, *Homo sapiens*: D= -0.2967, p-value<0.00001). Such gene flow resulted in higher autosomal affinity of *P. tapanuliensis* to *P. abelii* compared to *P. pygmaeus* in the PCA (Figure 2B), explaining the smaller amount of variance captured by PC2 (separating *P. tapanuliensis* from all other populations) compared to PC1 (separating *P. pygmaeus* from the Sumatran populations). The parameter estimates from a Bayesian full-likelihood analysis implemented in the software G-PhoCS were in good agreement with those obtained by the ABC analysis, although the split time between populations north and south of Lake Toba was more recent (~2.27 Ma, 95%-HPD: 2.21–2.35, Table S5). The G-PhoCS analysis revealed highly asymmetric gene flow between populations north and south of the Toba caldera, with much lower levels of gene flow into the Batang Toru population from the north than vice versa (Table S5).

The existence of two deep evolutionary lineages among extant Sumatran orangutans was corroborated by phylogenetic analyses based on whole mitochondrial genomes (Figure 4A), in which the deepest split occurred between populations north of Lake Toba and all other orangutans at ~3.97 Ma (95%-HPD: 2.35–5.57). Sumatran orangutans formed a paraphyletic group, with *P. tapanuliensis* being more closely related to the Bornean lineage from which it diverged ~2.41 Ma (1.26–3.42 Ma). In contrast, Bornean populations formed a monophyletic group with a very recent mitochondrial coalescence at ~160 ka (94–227 ka).

Due to strong female philopatry [28], gene flow in orangutans is almost exclusively male-mediated [29]. Consistent with these pronounced differences in dispersal behavior, phylogenetic analysis of extensive Y-chromosomal sequencing data revealed a comparatively recent coalescence of Y chromosomes of all extant orangutans ~430 ka (Figure 4B). The single available Y-haplotype from *P. tapanuliensis* was nested within the other Sumatran sequences, pointing at the occurrence of male-mediated gene flow across the Toba divide. Thus, in combination with our modeling results, the sex-specific data highlighted the impact of extraordinarily strong male-biased dispersal in the speciation process of orangutans.

Our analyses revealed significant divergence between *P. tapanuliensis* and *P. abelii* (Figs. 3B and 4A), and low levels of male-mediated gene flow (Figs. 3B and 4B), which, however, completely ceased 10–20 ka ago (Figure 3C). Populations north and south of Lake Toba on Sumatra had been in genetic contact for most of the time since their split, but there was a marked reduction in gene flow after ~100 ka (Figure 3C), consistent with habitat destruction caused by the Toba supereruption 73 ka ago [30]. However, *P. tapanuliensis* and *P. abelii* have been on independent evolutionary trajectories at least since the late Pleistocene/early Holocene, as gene flow between these populations has ceased completely 10–20 ka (Figure 3C) and is now impossible because of habitat loss in areas between the species' ranges [7].

Nowadays, most biologists would probably adopt an operational species definition such as: 'a species is a population (or group of populations) with fixed heritable differences from other such populations (or groups of populations)' [31]. With totally allopatric populations, a 'reproductive isolation' criterion, such as is still espoused by adherents of the biological species concept, is not possible [32, 33]. Notwithstanding a long-running debate about the role of gene flow during speciation and genetic interpretations of the species concept [34, 35], genomic studies have found evidence for many instances of recent or ongoing gene flow between taxa which are recognized as distinct and well-established species. This includes examples within each of the other three hominid genera. A recent genomic study using comparable methods to ours revealed extensive gene flow between *Gorilla gorilla* and *G. beringei* until ~20–30 ka [36]. Similar, albeit older and less extensive, admixture occurred between *Pan troglodytes* and *P. paniscus* [37], and was also reported for *Homo sapiens* and *H. neanderthalensis* [38]. *Pongo tapanuliensis* and *P. abelii* appear to be further examples, showing diagnostic phenotypic and other distinctions that had persisted in the past despite gene flow between them.

With a census size of fewer than 800 individuals [7], *P. tapanuliensis* is the least numerous of all great ape species [39]. Its range is located around 200 km from the closest population of *P. abelii* to the north (Figure 2A). A combination of small population size and geographic isolation is of particular high conservation concern, as it may lead to inbreeding depression [40] and threaten population persistence [41]. Highlighting this, we discovered extensive runs of homozygosity in the genomes of both *P. tapanuliensis* individuals (Figure S3), pointing at the occurrence of recent inbreeding.

To ensure long-term survival of *P. tapanuliensis*, conservation measures need to be implemented swiftly. Due to the rugged terrain, external threats have been primarily limited to road construction, illegal clearing of forests, hunting, killings during crop conflict and trade in orangutans [7, 11]. A hydro-electric development has been proposed recently in the area of highest orangutan density, which could impact up to 8% of *P. tapanuliensis'* habitat. This project might lead to further genetic

261 impoverishment and inbreeding, as it would jeopardize chances of maintaining habitat corridors
262 between the western and eastern range (Figure 1A), and smaller nature reserves, all of which maintain
263 small populations of *P. tapanuliensis.*

Author Contributions

Conceived the study and wrote the paper: MPMG, AlN, EM, MK, MGN, CG. Edited the manuscript: SW, GF, CvS, AS, TMB, DAM, TBS, TD, BG, FC, KSW, EV, POtW, PR, JB, MA, AnN. Carried out statistical analyses: MPMG, AlN, MGN, AnN, CG, MdM, TD, JA, MDR, AL, MP, JPM, MK, EM, AS, TMB. Provided samples, and behavioral and ecological data: MGN, MPMG, AnN, AlN, GF, JA, AL, MDR, BG, EJV, KSW, IS, JP, DPF, PR, WB. Performed sequencing: MPMG, IGG, MG, CR

Acknowledgments

We thank the following institutions and organizations for supporting our research: Indonesian State Ministry for Research and Technology, Sabah Wildlife Department, Ministry of Environment and Forestry of the Republic of Indonesia, Indonesian Institute of Sciences, Leuser International Foundation, Gunung Leuser National Park, Borneo Orangutan Survival Foundation, Agisoft, NVIDIA, and the 10 museums where we measured the specimens. This work was financially supported by University of Zurich (UZH) Forschungskredit grants FK-10 (MPMG), FK-15-103 (AlN), and FK-14-094 (TBS), Swiss National Science Foundation grant 3100A-116848 (MK, CvS), Leakey Foundation (MPMG), A.H. Schultz Foundation grants (MK, MPMG), UZH Research Priority Program 'Evolution in Action' (MK), the Arcus Foundation (EM), Australian National University (ANU) research fund (AnN), ANU Vice Chancellor Travel Grant (AnN), Australia Awards Scholarship-DFAT (AnN), ERC Starting Grant 260372 (TMB), EMBO YIP 2013 (TMB), MINECO BFU2014-55090-P, BFU2015-7116-ERC, BFU2015-6215-ERCU01, and MH106874 (TMB), Fundacio Zoo Barcelona (TMB), Julius–Klaus Foundation (MK), MINECO/FEDER BFU2016-77961-P (JB, MP), Gates Cambridge Trust (TD), and the Department of Anthropology at the University of Zurich. Novel raw sequencing data have been deposited into the European Nucleotide Archive (ENA; http://www.ebi.ac.uk/ena) under study accession number PRJEB19688.

References

1. Groves, C.P. (2001). Primate taxonomy, (Washington, D.C. ; London: Smithsonian Institution Press).
2. Xu, X., and Arnason, U. (1996). The mitochondrial DNA molecule of Sumatran orangutan and a molecular proposal for two (Bornean and Sumatran) species of orangutan. J. Mol. Evol. *43*, 431-437.
3. Wich, S.A., Utami Atmoko, S.S., Mitra Setia, T., and van Schaik, C.P. (2009). Orangutans: geographic variation in behavioral ecology and conservation, (Oxford University Press).
4. Nederlandsch-Indische Vereeniging tot Natuurbescherming (1939). Natuur in Zuid- en Oost-Borneo. Fauna, flora en natuurbescherming in de Zuider- en Ooster-Afdeeling van Borneo. In 3 Jaren Indisch natuur leven. Opstellen over landschappen, dieren en planten, tevens elfde verslag (1936-1938), Nederlandsch-Indische Vereeniging tot Natuurbescherming, ed. (Batavia, Indonesia), pp. 334-411.
5. Groves, C.P., Westwood, C., and Shea, B.T. (1992). Unfinished business - Mahalanobis and a clockwork orang. J. Hum. Evol. *22*, 327-340.
6. Groves, C.P. (1986). Systematics of the great apes. In Comparative primate biology, Vol.1: Systematics, evolution, and anatomy, D.R. Swindler and J. Erwin, eds. (New York: Alan R. Liss), pp. 187–217.
7. Wich, S.A., Singleton, I., Nowak, M.G., Utami Atmoko, S.S., Nisam, G., Arif, S.M., Putra, R.H., Ardi, R., Fredriksson, G., Usher, G., et al. (2016). Land-cover changes predict steep declines for the Sumatran orangutan (*Pongo abelii*). Sci. Adv. *2*, e1500789.
8. Laumonier, Y., Uryu, Y., Stüwe, M., Budiman, A., Setiabudi, B., and Hadian, O. (2010). Eco-floristic sectors and deforestation threats in Sumatra: identifying new conservation area network priorities for ecosystem-based land use planning. Biodivers. Conserv. *19*, 1153-1174.
9. Wich, S.A., Usher, G., Peters, H.H., Khakim, M.F.R., Nowak, M.G., and Fredriksson, G.M. (2014). Preliminary data on the highland Sumatran orangutans (*Pongo abelii*) of Batang Toru. In High Altitude Primates, B.N. Grow, S. Gursky-Doyen and A. Krzton, eds. (New York, NY: Springer New York), pp. 265-283.
10. Meijaard, E. (1997). A survey of some forested areas in South and Central Tapanuli, North Sumatra; new chances for orangutan conservation. (Wageningen: Tropenbos and the Golden Ark).
11. Wich, S.A., Fredriksson, G.M., Usher, G., Peters, H.H., Priatna, D., Basalamah, F., Susanto, W., and Kuhl, H. (2012). Hunting of Sumatran orang-utans and its importance in determining distribution and density. Biol. Conserv. *146*, 163-169.
12. Kramm, W. (1879). Tochtjes in Tapanoeli. Sumatra-Courant *20*, 1-2.
13. Miller, G.S. (1903). Mammals collected by Dr. W.L. Abbott on the coast and islands of northwest Sumatra. Proceedings US National Museum, Washington *26*, 437-484.
14. Shapiro, J.S. (1995). Morphometric variation in the orang utan (*Pongo pygmaeus*), with a comparison of inter- and intraspecific variability in the African apes. Volume PhD Dissertation. (Columbia University).
15. Uchida, A. (1998). Variation in tooth morphology of *Pongo pygmaeus*. J. Hum. Evol. *34*, 71-79.
16. Coolidge, H.J. (1933). *Pan paniscus*. Pigmy chimpanzee from south of the Congo river. Am. J. Phys. Anthropol. *18*, 1-59.
17. Shea, B.T., and Coolidge, H.J. (1988). Craniometric differentiation and systematics in the genus *Pan*. J. Hum. Evol. *17*, 671-685.
18. Groves, C.P. (2002). A history of gorilla taxonomy. In Gorilla biology: a multidisciplinary perspective, A.B. Taylor and M.L. Goldsmith, eds. (Cambridge: Cambridge University Press), pp. 15-34.
19. Stumpf, R.M., Polk, J.D., Oates, J.F., Jungers, W.L., Heesy, C.P., Groves, C.P., and Fleagle, J.G. (2002). Patterns of diversity in gorilla cranial morphology. In Gorilla biology: a multidisciplinary perspective, A.B. Taylor and M.L. Goldsmith, eds. (Cambridge: Cambridge University Press), pp. 35-61.

20. Hedrick, P.W. (2005). A standardized genetic differentiation measure. Evolution *59*, 1633-1638.
21. Jost, L. (2008). Gst and its relatives do not measure differentiation. Mol. Ecol. *17*, 4015-4026.
22. Whitlock, M.C. (2011). G'st and D do not replace Fst. Mol. Ecol. *20*, 1083-1091.
23. Nater, A., Arora, N., Greminger, M.P., van Schaik, C.P., Singleton, I., Wich, S.A., Fredriksson, G., Perwitasari-Farajallah, D., Pamungkas, J., and Krützen, M. (2013). Marked population structure and recent migration in the critically endangered Sumatran orangutan (*Pongo abelii*). J. Hered. *104*, 2-13.
24. Nielsen, R., and Wakeley, J. (2001). Distinguishing migration from isolation: a Markov chain Monte Carlo approach. Genetics *158*, 885-896.
25. Palsboll, P.J., Berube, M., Aguilar, A., Notarbartolo-Di-Sciara, G., and Nielsen, R. (2004). Discerning between recurrent gene flow and recent divergence under a finite-site mutation model applied to North Atlantic and Mediterranean Sea fin whale (Balaenoptera physalus) populations. Evolution *58*, 670-675.
26. Beaumont, M.A., Zhang, W.Y., and Balding, D.J. (2002). Approximate Bayesian computation in population genetics. Genetics *162*, 2025-2035.
27. Meijaard, E. (2004). Solving mammalian riddles: a reconstruction of the Tertiary and Quaternary distribution of mammals and their palaeoenvironments in island South-East Asia. (Australian National University), p. 2 v.
28. Arora, N., Van Noordwijk, M.A., Ackermann, C., Willems, E.P., Nater, A., Greminger, M., Nietlisbach, P., Dunkel, L.P., Utami Atmoko, S.S., Pamungkas, J., et al. (2012). Parentage-based pedigree reconstruction reveals female matrilineal clusters and male-biased dispersal in nongregarious Asian great apes, the Bornean orang-utans (*Pongo pygmaeus*). Molecular ecology *21*, 3352-3362.
29. Nater, A., Nietlisbach, P., Arora, N., van Schaik, C.P., van Noordwijk, M.A., Willems, E.P., Singleton, I., Wich, S.A., Goossens, B., Warren, K.S., et al. (2011). Sex-biased dispersal and volcanic activities shaped phylogeographic patterns of extant orangutans (genus: *Pongo*). Mol. Biol. Evol. *28*, 2275-2288.
30. Chesner, C.A., Rose, W.I., Deino, A., Drake, R., and Westgate, J.A. (1991). Eruptive history of earths largest Quaternary caldera (Toba, Indonesia) clarified. Geology *19*, 200-203.
31. Groves, C.P., and Grubb, P. (2011). Ungulate taxonomy, (Baltimore, Md.: Johns Hopkins University Press).
32. Coyne, J.A., and Orr, H.A. (2004). Speciation, (Sunderland, MA: Sinauer Associates, Inc.).
33. Mayr, E. (1963). Animal species and evolution, (Cambridge,: Belknap Press of Harvard University Press).
34. Arnold, M.L. (2016). Divergence with Genetic Exchange, (Oxford, UK: Oxford University Press).
35. Reznick, D.N., and Ricklefs, R.E. (2009). Darwin's bridge between microevolution and macroevolution. Nature *457*, 837-842.
36. Scally, A., Dutheil, J.Y., Hillier, L.W., Jordan, G.E., Goodhead, I., Herrero, J., Hobolth, A., Lappalainen, T., Mailund, T., Marques-Bonet, T., et al. (2012). Insights into hominid evolution from the gorilla genome sequence. Nature *483*, 169-175.
37. de Manuel, M., Kuhlwilm, M., Frandsen, P., Sousa, V.C., Desai, T., Prado-Martinez, J., Hernandez-Rodriguez, J., Dupanloup, I., Lao, O., Hallast, P., et al. (2016). Chimpanzee genomic diversity reveals ancient admixture with bonobos. Science *354*, 477.
38. Kuhlwilm, M., Gronau, I., Hubisz, M.J., de Filippo, C., Prado-Martinez, J., Kircher, M., Fu, Q., Burbano, H.A., Lalueza-Fox, C., de la Rasilla, M., et al. (2016). Ancient gene flow from early modern humans into Eastern Neanderthals. Nature *530*, 429-433.
39. IUCN (2016). IUCN Red List of Threatened Species. Version 2016.2.
40. Hedrick, P.W., and Kalinowski, S.T. (2000). Inbreeding depression in conservation biology. Annu. Rev. Ecol. Syst. *31*, 139-162.
41. Allendorf, F.W., Luikart, G., and Aitken, S.N. (2013). Conservation and the genetics of populations, 2nd Edition, (Hoboken: John Wiley & Sons).

42. Locke, D.P., Hillier, L.W., Warren, W.C., Worley, K.C., Nazareth, L.V., Muzny, D.M., Yang, S.-P., Wang, Z., Chinwalla, A.T., Minx, P., et al. (2011). Comparative and demographic analysis of orang-utan genomes. Nature *469*, 529-533.
43. Prado-Martinez, J., Sudmant, P.H., Kidd, J.M., Li, H., Kelley, J.L., Lorente-Galdos, B., Veeramah, K.R., Woerner, A.E., O/'Connor, T.D., Santpere, G., et al. (2013). Great ape genetic diversity and population history. Nature *499*, 471-475.
44. Arora, N., Nater, A., van Schaik, C.P., Willems, E.P., van Noordwijk, M.A., Goossens, B., Morf, N., Bastian, M., Knott, C., Morrogh-Bernard, H., et al. (2010). Effects of Pleistocene glaciations and rivers on the population structure of Bornean orangutans (*Pongo pygmaeus*). Proceedings of the National Academy of Sciences *107*, 21376-21381.
45. Nater, A., Nietlisbach, P., Arora, N., van Schaik, C.P., van Noordwijk, M.A., Willems, E.P., Singleton, I., Wich, S.A., Goossens, B., Warren, K.S., et al. (2011). Sex-biased dispersal and volcanic activities shaped phylogeographic patterns of extant orangutans (genus: *Pongo*). Molecular Biology and Evolution *28*, 2275-2288.
46. van Noordwijk, M.A., Arora, N., Willems, E.P., Dunkel, L.P., Amda, R.N., Mardianah, N., Ackermann, C., Krützen, M., and van Schaik, C.P. (2012). Female philopatry and its social benefits among Bornean orangutans. Behavioral Ecology and Sociobiology *66*, 823-834.
47. Morrogh-Bernard, H.C., Morf, N.V., Chivers, D.J., and Krützen, M. (2011). Dispersal patterns of orang-utans (*Pongo* spp.) in a Bornean peat-swamp forest. International Journal of Primatology *32*, 362-376.
48. Nietlisbach, P., Arora, N., Nater, A., Goossens, B., Van Schaik, C.P., and Krützen, M. (2012). Heavily male-biased long-distance dispersal of orang-utans (genus: *Pongo*), as revealed by Y-chromosomal and mitochondrial genetic markers. Molecular ecology *21*, 3173-3186.
49. Nater, A., Greminger, M.P., Arora, N., van Schaik, C.P., Goossens, B., Singleton, I., Verschoor, E.J., Warren, K.S., and Krützen, M. (2015). Reconstructing the demographic history of orang-utans using Approximate Bayesian Computation. Molecular Ecology *24*, 310-327.
50. Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular biology and evolution *29*, 1969-1973.
51. Tamura, K., and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution *10*, 512-526.
52. Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. Nature Methods *9*, 772-772.
53. Röhrer-Ertl, O. (1988). Research history, nomenclature, and taxonomy of the orang-utan. In Orang-utan Biology, J. Schwartz, ed. (Oxford, UK: Oxford University Press), pp. 7-18.
54. Hooijer, D.A. (1948). Prehistoric teeth of man and of the orang utan from Central Sumatra, with notes on the fossil orang utan from Java and Southern China. Zool Meded Rijksmus Leiden *29*, 175 - 183.
55. Drawhorn, G.M. (1994). The systematics and Paleodemography of fossil Orangutans (Genus *Pongo*). (University of California).
56. Harrison, T., Jin, C., Zhang, Y., Wang, Y., and Zhu, M. (2014). Fossil Pongo from the Early Pleistocene Gigantopithecus fauna of Chongzuo, Guangxi, southern China. Quaternary International *354*, 59-67.
57. de Vos, J. (1983). The *Pongo* faunas from Java and Sumatra and their significance for biostratigraphical and paleo-ecological interpretations. Proceedings of the Koninklijke Akademie van Wetenschappen. Series B *86*, 417-425.
58. Bacon, A.-M., Westaway, K., Antoine, P.-O., Duringer, P., Blin, A., Demeter, F., Ponche, J.-L., Zhao, J.-X., Barnes, L.M., Sayavonkhamdy, T., et al. (2015). Late Pleistocene mammalian assemblages of Southeast Asia: New dating, mortality profiles and evolution of the predator–prey relationships in an environmental context. Palaeogeography, Palaeoclimatology, Palaeoecology *422*, 101-127.
59. Louys, J. (2012). Mammal community structure of Sundanese fossil assemblages from the Late Pleistocene, and a discussion on the ecological effects of the Toba eruption. Quaternary International *258*, 80-87.

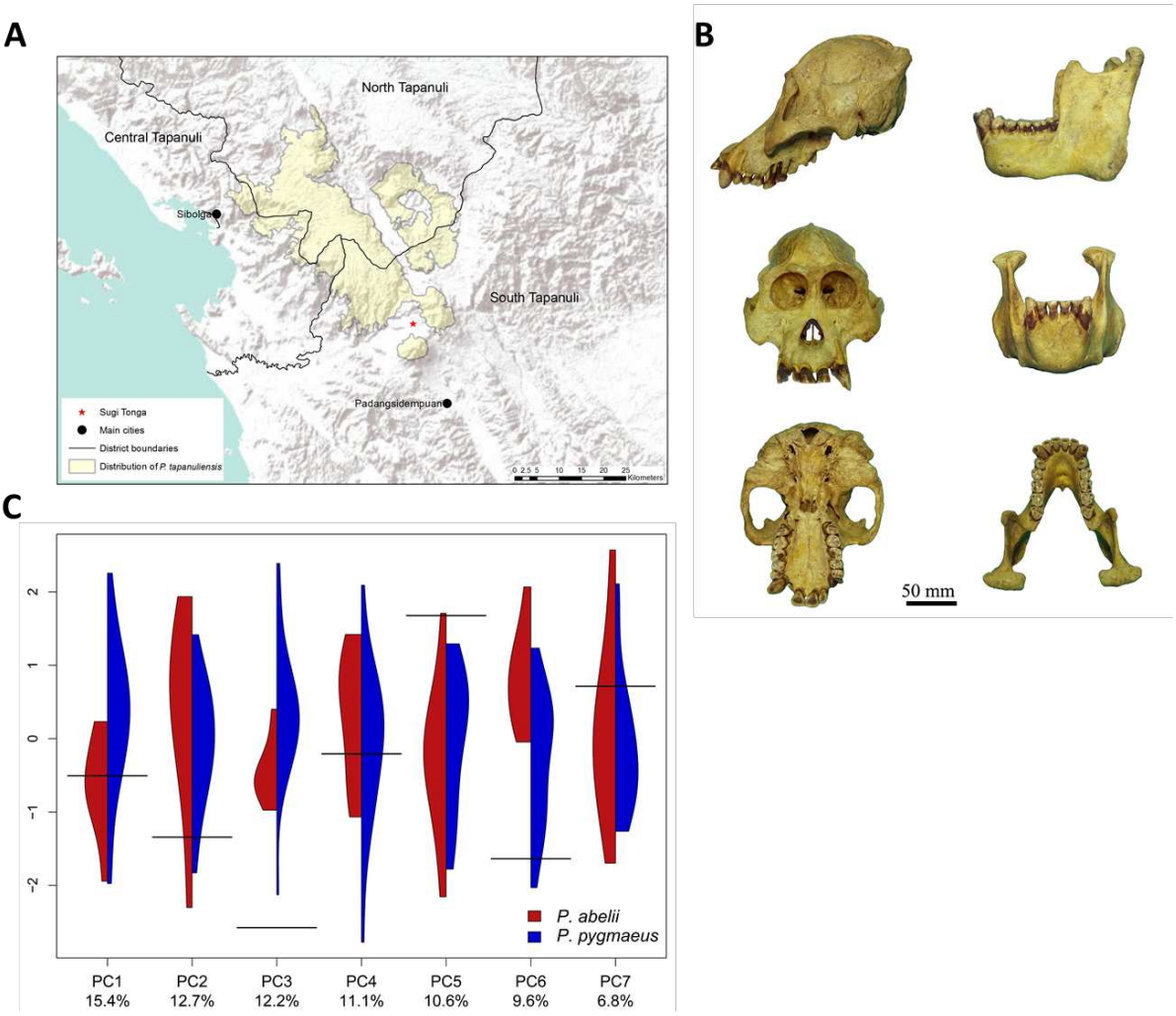
60. Schwartz, J.H., Vu The, L., Nguyen Lan, C., Le Trung, K., and Tattersall, I. (1995). A review of the Pleistocene hominoid fauna of the Socialist Republic of Vietnam (excluding Hylobatidae).
61. Plavcan, J.M. (1994). Comparison of four simple methods for estimating sexual dimorphism in fossils. Am J Phys Anthropol *94*, 465-476.
62. Greminger, M.P., Stolting, K., Nater, A., Goossens, B., Arora, N., Bruggmann, R., Patrignani, A., Nussberger, B., Sharma, R., Kraus, R.H., et al. (2014). Generation of SNP datasets for orangutan population genomics using improved reduced-representation sequencing and direct comparisons of SNP calling algorithms. BMC genomics *15*, 16.
63. Andrews, S. (2012). FastQC. A quality control tool for high throughput sequence data.
64. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics *25*, 1754-1760.
65. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research *20*, 1297-1303.
66. DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet *43*, 491-498.
67. Derrien, T., Estellé, J., Marco Sola, S., Knowles, D.G., Raineri, E., Guigó, R., and Ribeca, P. (2012). Fast Computation and Applications of Genome Mappability. PLoS ONE *7*, e30377.
68. Auton, A., and McVean, G. (2007). Recombination rate estimation in the presence of hotspots. Genome Research *17*, 1219-1227.
69. Auton, A., Fledel-Alon, A., Pfeifer, S., Venn, O., Segurel, L., Street, T., Leffler, E.M., Bowden, R., Aneas, I., Broxholme, J., et al. (2012). A fine-scale chimpanzee genetic map from population sequencing. Science *336*, 193-198.
70. Delaneau, O., Marchini, J., and Zagury, J.F. (2012). A linear complexity phasing method for thousands of genomes. Nat Methods *9*, 179-181.
71. Delaneau, O., Howie, B., Cox, A.J., Zagury, J.F., and Marchini, J. (2013). Haplotype estimation using sequencing reads. American Journal of Human Genetics *93*, 687-696.
72. McQuillan, R., Leutenegger, A.L., Abdel-Rahman, R., Franklin, C.S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A., et al. (2008). Runs of homozygosity in European populations. American Journal of Human Genetics *83*, 359-372.
73. Pemberton, Trevor J., Absher, D., Feldman, Marcus W., Myers, Richard M., Rosenberg, Noah A., and Li, Jun Z. (2012). Genomic Patterns of Homozygosity in Worldwide Human Populations. The American Journal of Human Genetics *91*, 275-292.
74. Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In Nucleic acids symposium series, Volume 41. pp. 95-98.
75. Roos, C., Zinner, D., Kubatko, L., Schwarz, C., Yang, M., Meyer, D., Nash, S., Xing, J., Batzer, M., Brameier, M., et al. (2011). Nuclear versus mitochondrial DNA: evidence for hybridization in colobine monkeys. BMC Evolutionary Biology *11*, 77.
76. Thalmann, O., Serre, D., Hofreiter, M., Lukas, D., Eriksson, J., and Vigilant, L. (2005). Nuclear insertions help and hinder inference of the evolutionary history of gorilla mtDNA. Molecular Ecology *14*, 179-188.
77. Steiper, M.E., and Young, N.M. (2006). Primate molecular divergence dates. Molecular phylogenetics and evolution *41*, 384-394.
78. Bellott, D.W., Hughes, J.F., Skaletsky, H., Brown, L.G., Pyntikova, T., Cho, T.-J., Koutseva, N., Zaghlul, S., Graves, T., and Rock, S. (2014). Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. Nature *508*, 494-499.
79. Soh, Y.S., Alföldi, J., Pyntikova, T., Brown, L.G., Graves, T., Minx, P.J., Fulton, R.S., Kremitzki, C., Koutseva, N., and Mueller, J.L. (2014). Sequencing the mouse Y chromosome reveals convergent gene acquisition and amplification on both sex chromosomes. Cell *159*, 800-813.

80. Hughes, J.F., Skaletsky, H., Pyntikova, T., Graves, T.A., van Daalen, S.K., Minx, P.J., Fulton, R.S., McGrath, S.D., Locke, D.P., and Friedman, C. (2010). Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content. Nature *463*, 536-539.
81. Wei, W., Ayub, Q., Chen, Y., McCarthy, S., Hou, Y., Carbone, I., Xue, Y., and Tyler-Smith, C. (2013). A calibrated human Y-chromosomal phylogeny based on resequencing. Genome research *23*, 388-395.
82. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Subgroup, G.P.D.P. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.
83. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156-2158.
84. Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. In Lectures on Mathematics in the Life Sciences, Volume 17. pp. 57-86.
85. Posada, D. (2003). Using MODELTEST and PAUP* to Select a Model of Nucleotide Substitution. In Current Protocols in Bioinformatics. (John Wiley & Sons, Inc.).
86. Drummond, A.J., Ho, S.Y., Phillips, M.J., and Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. PLoS biology *4*, e88.
87. Yang, Z., and Rannala, B. (2006). Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. Molecular biology and evolution *23*, 212-226.
88. Brunet, M., Guy, F., Pilbeam, D., Mackaye, H.T., Likius, A., Ahounta, D., Beauvilain, A., Blondel, C., Bocherens, H., and Boisserie, J.-R. (2002). A new hominid from the Upper Miocene of Chad, Central Africa. Nature *418*, 145-151.
89. Vignaud, P., Duringer, P., Mackaye, H.T., Likius, A., Blondel, C., Boisserie, J.-R., De Bonis, L., Eisenmann, V., Etienne, M.-E., and Geraads, D. (2002). Geology and palaeontology of the Upper Miocene Toros-Menalla hominid locality, Chad. Nature *418*, 152-155.
90. Raaum, R.L., Sterner, K.N., Noviello, C.M., Stewart, C.-B., and Disotell, T.R. (2005). Catarrhine primate divergence dates estimated from complete mitochondrial genomes: concordance with fossil and nuclear DNA evidence. J Hum Evol *48*, 237-257.
91. Rambaut, A., Suchard, M.A., Xie, D., and Drummond, A.J. (2014). Tracer v1.6.
92. Rambaut, A. (2012). FigTree version 1.4.
93. Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular biology and evolution *30*, 2725-2729.
94. Scally, A., and Durbin, R. (2012). Revising the human mutation rate: implications for understanding human evolution. Nature Reviews Genetics *13*, 745-753.
95. Ségurel, L., Wyman, M.J., and Przeworski, M. (2014). Determinants of Mutation Rate Variation in the Human Germline. Annual Review of Genomics and Human Genetics *15*, 47-70.
96. Venn, O., Turner, I., Mathieson, I., de Groot, N., Bontrop, R., and McVean, G. (2014). Strong male bias drives germline mutation in chimpanzees. Science *344*, 1272-1275.
97. Lipson, M., Loh, P.-R., Sankararaman, S., Patterson, N., Berger, B., and Reich, D. (2015). Calibrating the human mutation rate via ancestral recombination density in diploid genomes. PLoS Genet *11*, e1005550.
98. Carbone, L., Alan Harris, R., Gnerre, S., Veeramah, K.R., Lorente-Galdos, B., Huddleston, J., Meyer, T.J., Herrero, J., Roos, C., Aken, B., et al. (2014). Gibbon genome and the fast karyotype evolution of small apes. Nature *513*, 195-201.
99. Wich, S., De Vries, H., Ancrenaz, M., Perkins, L., Shumaker, R., Suzuki, A., and Van Schaik, C. (2009). Orangutan life history variation. In Orangutans - Geographic Variation in Behavioral Ecology and Conservation S.A. Wich, S.S. Utami Atmoko, T. Mitra Setia and C.P. van Schaik, eds. (Oxford University Press), pp. 65-75.
100. Team, R.D.C. (2010). R: a language and environment for statistical computing. (Vienna, Austria: R Foundation for Statistical Computing).

101. Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Research *19*, 1655-1664.
102. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet *81*, 559-575.
103. Schiffels, S., and Durbin, R. (2014). Inferring human population size and separation history from multiple genome sequences. Nat. Genet. *46*, 919-925.
104. Robinson, J.D., Bunnefeld, L., Hearn, J., Stone, G.N., and Hickerson, M.J. (2014). ABC inference of multi-population divergence with admixture from unphased population genomic data. Mol. Ecol. *23*, 4458-4471.
105. Excoffier, L., Smouse, P.E., and Quattro, J.M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes - application to human mitochondrial DNA restriction data. Genetics *131*, 479-491.
106. Hudson, R.R. (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. Bioinformatics *18*, 337-338.
107. Le Cao, K.A., Gonzalez, I., and Dejean, S. (2009). integrOmics: an R package to unravel relationships between two omics datasets. Bioinformatics *25*, 2855-2856.
108. Csillery, K., Francois, O., and Blum, M.G.B. (2012). abc: an R package for approximate Bayesian computation (ABC). Methods Ecol. Evol. *3*, 475-479.
109. Mevik, B.H., and Wehrens, R. (2007). The pls package: principal component and partial least squares regression in R. J. Stat. Softw. *18*.
110. Wegmann, D., Leuenberger, C., and Excoffier, L. (2009). Efficient Approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. Genetics *182*, 1207-1218.
111. Leuenberger, C., and Wegmann, D. (2010). Bayesian computation and model selection without likelihoods. Genetics *184*, 243-252.
112. Wegmann, D., Leuenberger, C., Neuenschwander, S., and Excoffier, L. (2010). ABCtoolbox: a versatile toolkit for approximate Bayesian computations. BMC Bioinformatics *11*.
113. Cook, S.R., Gelman, A., and Rubin, D.B. (2006). Validation of software for Bayesian models using posterior quantiles. J. Comput. Graph. Stat. *15*, 675-692.
114. Rice, W.R. (1989). Analyzing tables of statistical tests. Evolution *43*, 223-225.
115. Gronau, I., Hubisz, M.J., Gulko, B., Danko, C.G., and Siepel, A. (2011). Bayesian inference of ancient human demography from individual genome sequences. Nat. Genet. *43*, 1031-1034.
116. Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M.A., and Alekseyenko, A.V. (2012). Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Mol Biol Evol *29*, 2157-2167.
117. Raftery, A.E., Newton, M.A., Satagopan, J.M., and Krivitsky, P.N. (2007). Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. In Bayesian Statistics, J.M. Bernardo, M.J. Bayarri and J.O. Berger, eds. (Oxford: Oxford University Press), pp. 1-45.
118. Röhrer-Ertl, O. (1984). Orang-utan Studien, (Neuried, Germany: Hieronymus Verlag).
119. Röhrer-Ertl, O. (1988). Cranial growth. In Orang-utan Biology, J. Schwartz, ed. (Oxford, UK: Oxford University Press), pp. 201-224.
120. Courtenay, J., Groves, C., and Andrews, P. (1988). Inter- or intra-island variation? An assessment of the differences between Bornean and Sumatran orang-utans. In Orang-utan biology, H. Schwartz, ed. (Oxford, England: Oxford University Press), pp. 19-29.
121. Uchida, A. (1998). Variation in tooth morphology of *Pongo pygmaeus.* J Hum Evol *34*, 71-79.
122. Taylor, A.B. (2006). Feeding behavior, diet, and the functional consequences of jaw form in orangutans, with implications for the evolution of Pongo. J Hum Evol *50*, 377-393.
123. Taylor, A.B. (2009). The functional significance of variation in jaw form in orangutans. In Orangutans: geographic variation in behavioral ecology and conservation, S.A. Wich, S.U. Atmoko, T.M. Setia and C.P. van Schaik, eds. (Oxford, UK.: Oxford University Press), pp. 15-31.

124. Tukey, J.W. (1977). Exploratory data analysis, (London, UK: Addison-Wesley Publishing Company).
125. Tabachnick, B.G., and Fidell, L.S. (2013). Using multivariate statistics, 6th ed, (New York, USA: Pearson).
126. R Core Development Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing. http://www.R-project.org/. (Vienna, Austria).
127. Kaiser, H.F. (1960). The application of electronic computers to factor analysis. Education and Psychological Measurement *20*, 141-151.
128. Revelle, W. (2016). Psych: procedures for personality and psychological research. http://CRAN.R-project.org/package=psych Version =1.6.4, (Evanston, Illinois, USA: Northwestern University).
129. Davila-Ross, M. (2004). The long calls of wild male orangutans: A phylogenetic approach. Volume PhD. (Hannover, Germany: Institut für Zoologie, Tierärztliche Hochshule Hannover).
130. Davila-Ross, M., and Geissmann, T. (2007). Call diversity of wild male orangutans: a phylogenetic approach. Am. J. Primatol. *69*, 305-324.
131. Lameira, A.R., and Wich, S.A. (2008). Orangutan Long Call Degradation and Individuality Over Distance: A Playback Approach. Int. J. Primatol. *29*, 615-625.
132. Delgado, R.A., Lameira, A.R., Davila Ross, M., Husson, S.J., Morrogh-Bernard, H.C., and Wich, S.A. (2009). Geographical variation in orangutan long calls. In Orangutans: Geographic variation in behavioral ecology and conservation, S.A. Wich, S.S. Utami Atmoko, T. Mitra Setia and C.P. van Schaik, eds. (Oxford, UK: Oxford University Press), pp. 215-224.
133. Darul Sukma, W.P., Dai, J., Hidayat, A., Yayat, A.H., Sumulyadi, H.Y., Hendra, S., Buurman, P., and Balsem, T. (1990). Explanatory booklet of the land unit and soil map of the Sidikalang sheet (618), Sumatra. (Bogor, Indonesia: Centre for Soil and Agroclimate Research).
134. Darul Sukma, W.P., Suratman, Hidayat, J.A., and Budhi, P.G. (1990). Explanatory booklet of the land unit and soil map of the Tapaktuan sheet (519), Sumatra, (Bogor, Indonesia: Centre for Soil and Agroclimate Research).
135. Darul Sukma, W.P., Suratman, Hidayat, J.A., and Budi, P.G. (1990). Explanatory booklet of the land unit and soil map of the Lho'Kruet sheet (420), Sumatra, (Bogor, Indonesia: Centre for Soil and Agroclimate Research).
136. Darul Sukma, W.P., Verhagen, V., Dai, J., Buurman, P., Balsem, T., Suratman, and Vejre, H. (1990). Explanatory booklet of the land unit and soil map of the Takengon sheet (520), Sumatra, (Bogor, Indonesia: Centre for Soil and Agroclimate Research).
137. Hidayat, A., Verhagen, A., Darul Sukma, W.P., Buurman, P., Balsem, T., Suratman, and Vejre, H. (1990). Explanatory booklet of the land unit and soil map of the Lhokseumawe (521) and Simpangulim (621) sheets, Sumatra, (Centre for Soil and Agroclimate Research).
138. Hikmatullah, Wahyunto, Chendy, T.F., Dai, J., and Hidayat, A. (1990). Explanatory booklet of the land unit and soil map of the Langsa (620) sheet, Sumatra, (Bogor, Indonesia: Centre for Soil and Agroclimate Research).
139. Subardja, D., Djuanda, K., Hadian, Y., Samdan, C.D., Mulyadi, Y., Supriatna, W., and Dai, J. (1990). Explanatory booklet of the land unit and soil map of the Sibolga (617) and Padangsidempuan (717) sheets, Sumatra, (Bogor, Indonesia: Centre for Soil and Agroclimate Research).
140. Wahyunto, Puksi, D.S., Rochman, A., Wahdini, W., Paidi, Dai, J., Hidayat, A., Buurman, P., and Balsem, T. (1990). Explanatory booklet of the land unit and soil map of the Medan (619) sheet, Sumatra, (Bogor, Indonesia: Centre for Soil and Agroclimate Research).
141. Hall, R., van Hattum, M.W.A., and Spakman, W. (2008). Impact of India–Asia collision on SE Asia: The record in Borneo. Tectonophysics *451*, 366-389.
142. Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology *25*, 1965-1978.

143. Wich, S.A., Singleton, I., Nowak, M.G., Utami Atmoko, S.S., Nisam, G., Arif, S.M., Putra, R.H., Ardi, R., Fredriksson, G., Usher, G., et al. (2016). Land-cover changes predict steep declines for the Sumatran orangutan (*Pongo abelii*). Science Advances *2*, e1500789.
144. Wich, S.A., Atmoko, S.U., Setia, T.M., and van Schaik, C. (2009). Orangutans. Geographic variation in behavioral ecology and conservation, (Oxford, UK: Oxford University Press).
145. Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic acids symposium series *41*, 95-98.
146. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078-2079.
147. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156-2158.
148. Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. Genetics *192*, 1065-1093.
149. Venables, W.N., and Ripley, B.D. (2002). Modern applied statistics with S, 4th edition, (New York, USA: Springer).



684

685 **Figure 1. Morphological evidence supporting a new orangutan species.** A) Current distribution of
686 *Pongo tapanuliensis* on Sumatra. The holotype locality is marked with a red star. The area shown in
687 the map is indicated in Figure 2A. B) Holotype skull and mandible of *P. tapanuliensis* from a recently
688 deceased individual from Batang Toru. See also Figure S1, Tables S1 and S2. C) Violin plots of the
689 first seven principal components of 26 cranio-mandibular morphological variables of 8 north
690 Sumatran *P. abelii* and 19 Bornean *P. pygmaeus* individuals of similar developmental state as the
691 holotype skull (black horizontal lines). See also Figure S2.

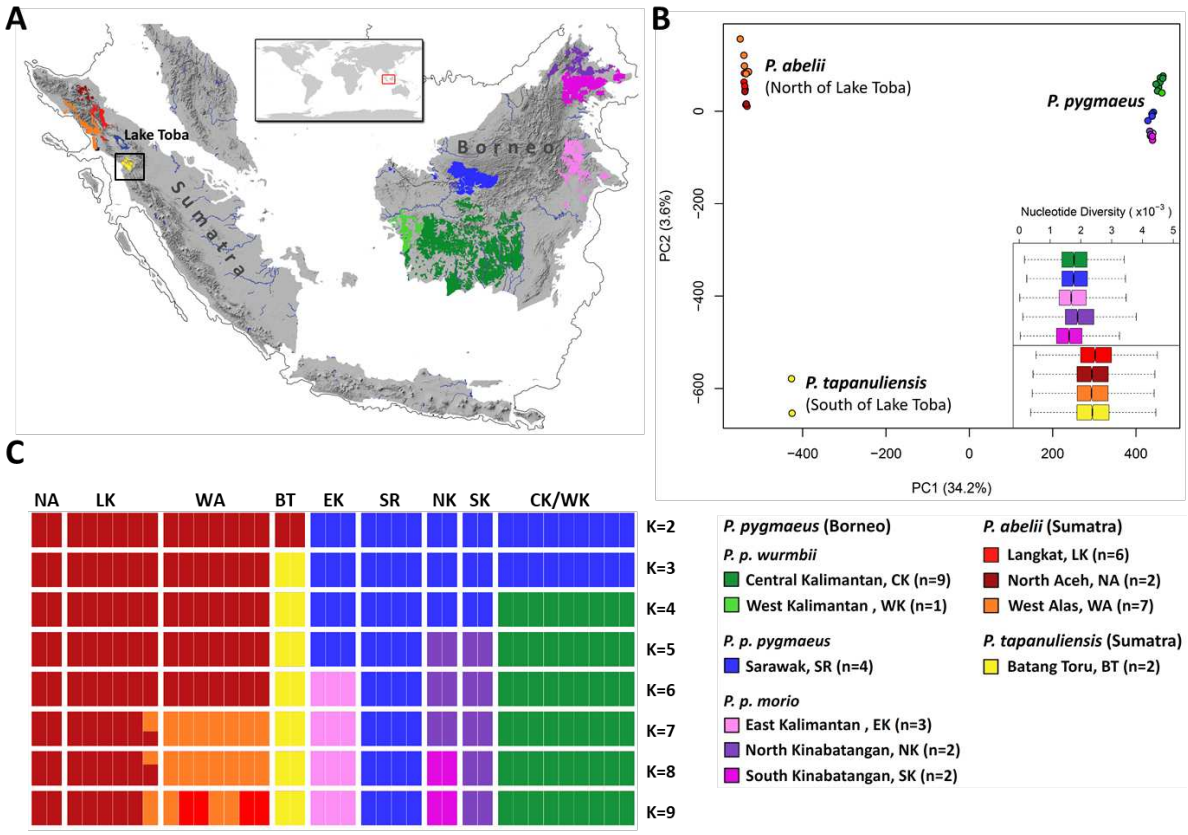
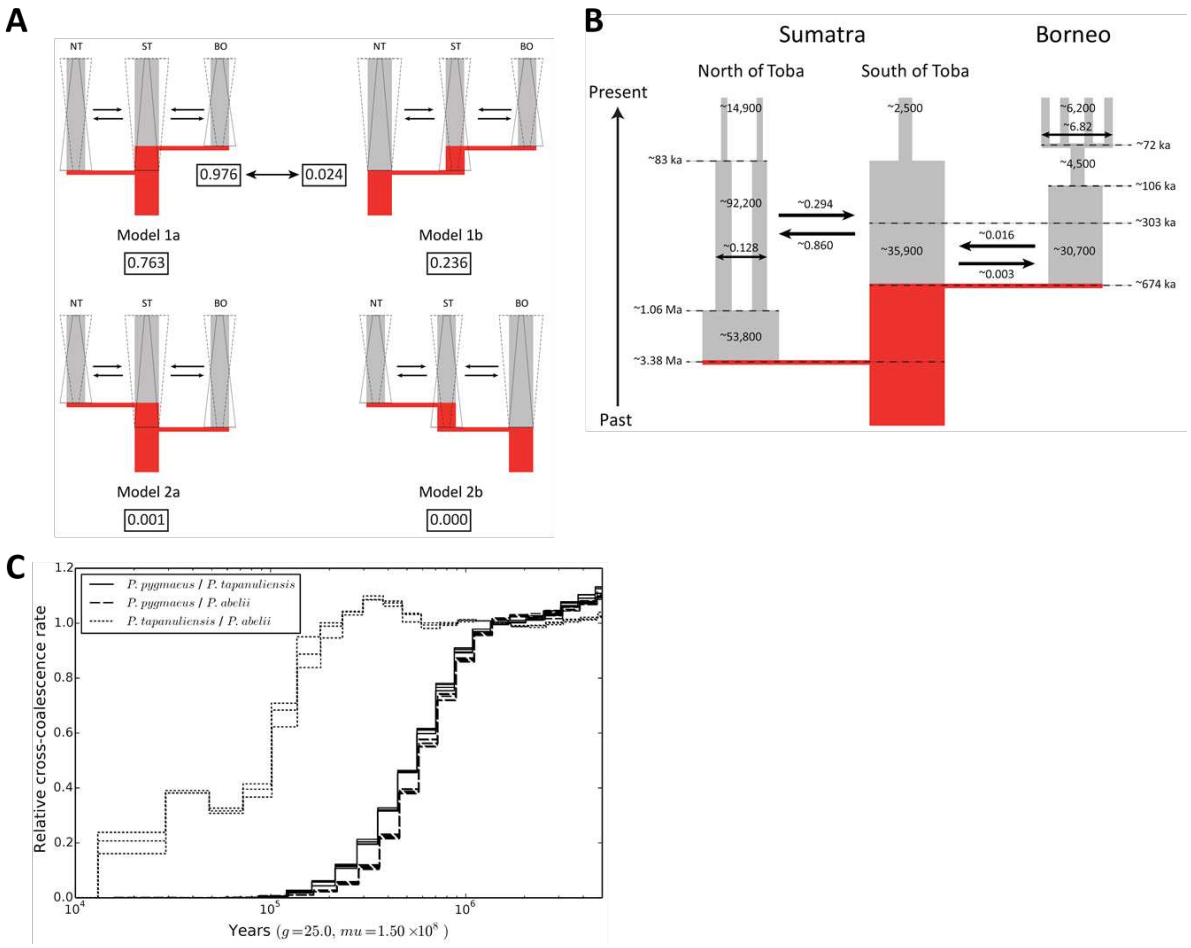


Figure 2. Distribution, genomic diversity, and population structure of the genus *Pongo*. A) Sampling areas across the current distribution of orangutans. The contour indicates the extent of the exposed Sunda Shelf during the last glacial maximum. The black rectangle delimits the area shown in Figure 1A. n = numbers of sequenced individuals. See also Table S4. B) Principal component analysis of genomic diversity in *Pongo*. Axis labels show the percentages of the total variance explained by the first two principal components. Colored bars in the insert represent the distribution of nucleotide diversity in genome-wide 1-Mb windows across sampling areas. C) Bayesian clustering analysis of population structure using the program ADMIXTURE. Each vertical bar depicts an individual, with colors representing the inferred ancestry proportions with different assumed numbers of genetic clusters (K, horizontal sections).



703

704 **Figure 3. Demographic history and gene flow in *Pongo*.** A) Model selection by Approximate
705 Bayesian Computation (ABC) of plausible colonization histories of orangutans on Sundaland. The
706 ABC analyses are based on the comparison of ~3,000 non-coding 2-kb loci randomly distributed
707 across the genome with corresponding data simulated under the different demographic models. The
708 numbers in the black boxes indicate the model's posterior probability. NT = Sumatran populations
709 north of Lake Toba, ST = the Sumatran population of Batang Toru south of Lake Toba, BO =
710 Bornean populations. B) ABC parameter estimates based on the full demographic model with
711 colonization pattern inferred in panel A. Numbers in grey rectangles represent point estimates of
712 effective population size (N_e). Arrows indicate gene flow among populations, numbers above the
713 arrows represent point estimates of numbers of migrants per generation. C) Relative cross-coalescent
714 rate (RCCR) analysis for between-species pairs of phased high-coverage genomes. A RCCR close to
715 1 indicates extensive gene flow between species, while a ratio close to 0 indicates genetic isolation
716 between species pairs. The x-axis shows time scaled in years, assuming a generation time of 25 years
717 and an autosomal mutation rate of 1.5×10^{-8} per site per generation. See also Figure S3.

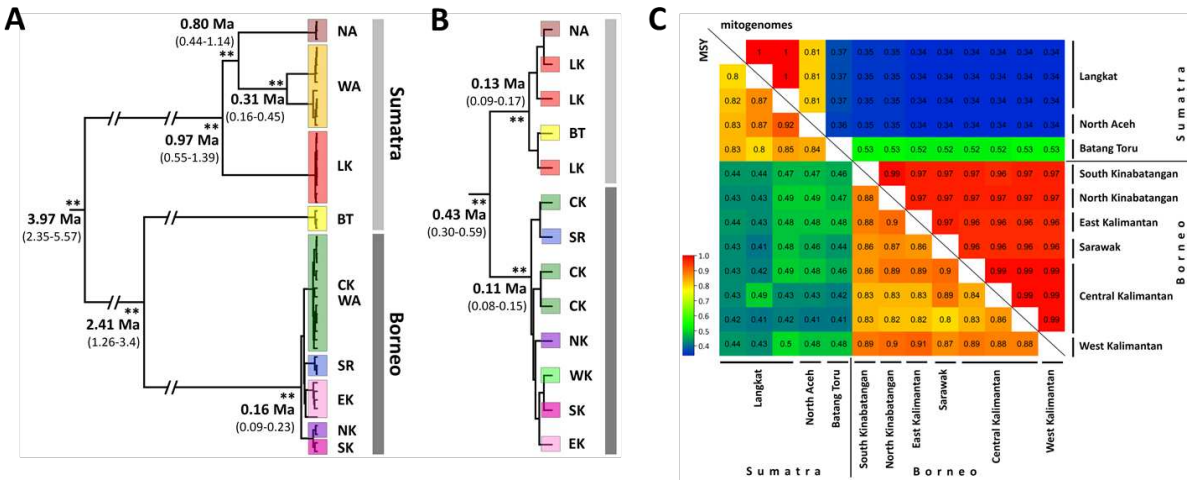


Figure 4. Sex-specific evolutionary history of orangutans. Bayesian phylogenetic trees for (A) mitochondrial genomes and (B) Y chromosomes. The mitochondrial tree is rooted with a human and a central chimpanzee sequence, the Y chromosome tree with a human sequence (not shown). ** Posterior probability = 1.00. C) Genotype-sharing matrix for mitogenomes (above the diagonal) and Y chromosomes (below the diagonal) for all analyzed male orangutans. A value of 1 indicates that two males have identical genotypes at all polymorphic sites; a value of 0 means that they have different genotypes at all variable positions.

CONTACT FOR RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Michael Krützen (michael.kruetzen@aim.uzh.ch).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Sample collection and population assignment for genomic analysis

Our sample set comprised genomes from 37 orangutans, representing the entire geographic range of extant orangutans (Figure 2A). We obtained whole-genome sequencing data for the study individuals from three different sources (Table S4): (i) genomes of 17 orangutans were sequenced for this study. Data for 20 individuals were obtained from (ii) Locke *et al.* [42] (n=10) and (iii) Prado-Martinez *et al.* [43] (n=10). All individuals were wild-born, except for five orangutans which were first-generation offspring of wild-born parents of the same species (Table S4).

Population provenance of the previously sequenced orangutans [42, 43] was largely unknown. We identified their most likely natal area based on mtDNA haplotype clustering in a phylogenetic tree together with samples of known geographic provenance. Because of extreme female philopatry in orangutans, mtDNA haplotypes are reliable indicators for the population of origin [28, 44-48]. Using three concatenated mtDNA genes (16S ribosomal DNA, Cytochrome b, and NADH-ubiquinone oxidoreductase chain 3), we constructed a Bayesian tree, including 127 non-invasively sampled wild orangutans from 15 geographic regions representing all known extant orangutan populations [45, 49]. Gene sequences of our study individuals were extracted from their complete mitochondrial genome sequences. The phylogenetic tree was built with BEAST v1.8.0. [50], as described in Nater *et al.* [45], applying a TN93+I substitution model [51] as determined by jModelTest v2.1.4. [52].

Using the mitochondrial tree, we assigned all previously sequenced orangutans [42, 43] to their most likely population of origin. Our sample assignment revealed incomplete geographic representation of the genus *Pongo* in previous studies. To achieve a more complete representation of extant orangutans, we sequenced genomes of 17 wild-born orangutans mainly from areas with little or no previous sample coverage. Detailed provenance information for these individuals is provided in Table S4.

Samples for morphological analysis

We conducted comparative morphological analyses of 34 adult male orangutans from 10 institutions housing osteological specimens. A single adult male skeleton from the Batang Toru population was available for study, having died from injuries sustained in an orangutan-human conflict situation in November 2013. To account for potential morphological differences related to developmental stage [14, 53], our analyses included only males at a similar developmental stage as the Batang Toru

758 specimen, *i.e.*, having a sagittal crest of <10 mm in height. In addition to the single available Batang
759 Toru male, our extant sample comprises specimens from the two currently recognised species, the
760 north Sumatran *Pongo abelii* (n=8) and the Bornean *P. pygmaeus* (n=25).

761 We also evaluated the relationship of the dental material between the Batang Toru specimen and those
762 of the Late Pleistocene fossil material found within the Djamboe, Lida Ajer, and Sibrambang caves
763 near Padang, Sumatra, all of which has been previously described by Hooijer [54]. Some scholars
764 have suggested that the fossil material may represent multiple species [55, 56]. However, Hooijer had
765 more than adequately shown that the variation in dental morphology observed within the three cave
766 assemblages can easily be accommodated within a single species [54]. As only teeth were present in
767 the described cave material, many of which also have gnaw marks, taphonomic processes (*e.g.*,
768 porcupines as accumulating agents) are thought to have largely shaped the cave material [57, 58] and
769 thus may account for the appearance of size differences among the cave samples [55, 56].
770 Furthermore, the similarities in the reconstructed age of the cave material (~128-118 ka or ~80-60 ka
771 [57-59]), and the fact that the presence of more than one large-bodied ape species is an uncommon
772 feature in both fossil and extant Southeast Asian faunal assemblages [60], makes it highly unlikely
773 that multiple large-bodied ape species co-existed within the area at a given time. For purposes of
774 discussion here, we collectively refer to the Padang fossil material as *P. p. palaeosumatrensis*, as
775 described by Hooijer [54].

776 As the comparative fossil sample likely comprises various age-sex classes [54], we divided the fossil
777 sample into two portions above and below the mean for each respective tooth utilized in this study.
778 We considered samples above the mean to represent larger individuals, which we attribute to “males”,
779 and the ones below to being smaller individuals, which we attribute to “females” [61]. We only used
780 the “male” samples in comparison to our extant male comparative orangutan sample.

METHOD DETAILS

Whole-genome sequencing

To obtain sufficient amounts of DNA, we collected blood samples from confiscated orangutans at rehabilitation centres, including the Sumatran Orangutan Conservation Program (SOCP) in Medan, BOS Wanariset Orangutan Reintroduction Project in East Kalimantan, Semongok Wildlife Rehabilitation Centre in Sarawak, and Sepilok Orangutan Rehabilitation Centre in Sabah. We took whole blood samples during routine veterinary examinations and stored in EDTA blood collection tubes at -20°C. The collection and transport of samples were conducted in strict accordance with Indonesian, Malaysian and international regulations. Samples were transferred to Zurich under the Convention on International Trade of Endangered Species in Fauna and Flora (CITES) permit numbers 4872/2010 (Sabah), and 06968/IV/SATS-LN/2005 (Indonesia).

We extracted genomic DNA using the Gentra Puregene Blood Kit (Qiagen) but modified the protocol for clotted blood as described in Greminger *et al.* [62]. We sequenced individuals on two to three lanes on an Illumina HiSeq 2000 in paired end (2 x 101 bp) mode. Sample PP_5062 was sequenced at the Functional Genomics Center in Zurich (Switzerland), the other individuals at the Centre Nacional d'Anàlisi Genòmica in Barcelona (Spain), as the individuals of Prado-Martinez *et al.* [43]. On average, we generated ~1.1×10^9 raw Illumina reads per individual.

Read mapping

We followed identical bioinformatical procedures for all 37 study individuals, using the same software versions. We quality-checked raw Illumina sequencing reads with FastQC v0.10.1. [63] and mapped to the orangutan reference genome *ponAbe2* [42] using the Burrows-Wheeler Aligner (BWA-MEM) v0.7.5 [64] in paired-end mode with default read alignment penalty scores. We used Picard v1.101 (http://picard.sourceforge.net/) to add read groups, convert sequence alignment/map (SAM) files to binary alignment/map (BAM) files, merge BAM files for each individual, and to mark optical and PCR duplicates. We filtered out duplicated reads, bad read mates, reads with mapping quality zero, and reads that mapped ambiguously.

We performed local realignment around indels and empirical base quality score recalibration (BQSR) with the Genome Analysis Toolkit (GATK) v3.2.2. [65, 66]. The BQSR process empirically calculates more accurate base quality scores (*i.e.*, Phred-scaled probability of error) than those emitted by the sequencing machines through analysing the covariation among several characteristics of a base (e.g. position within the read, sequencing cycle, previous base, etc.) and its status of matching the reference sequence or not. To account for true sequence variation in the data set, the model requires a database of known polymorphic sites ('known sites') which are skipped over in the recalibration algorithm. Since no suitable set of 'known sites' was available for the complete genus *Pongo*, we

preliminary identified confident SNPs from our data. For this, we performed an initial round of SNP calling on unrecalibrated BAM files with the *UnifiedGenotyper* of the GATK. Single nucleotide polymorphisms were called separately for Bornean and Sumatran orangutans in multi-sample mode (*i.e.*, joint analysis of all individuals per island), creating two variant call (VCF) files. In addition, we produced a third VCF file jointly analysing all study individuals in order to capture genus-wide low frequency alleles. We applied the following hard quality filter criteria on all three VCF files: QUAL < 50.0 || QD < 2.0 || FS > 60.0 || MQ < 40.0 || HaplotypeScore > 13.0 || MappingQualityRankSum < -12.5 || ReadPosRankSum < -8.0. Additionally, we calculated the mean and standard deviation of sequencing depth over all samples and filtered all sites with a site-wise coverage more than five standard deviations above the mean. We merged the three hard filtered VCF files and took SNPs as 'known sites' for BQSR with the GATK. The walkers CountReads and DepthOfCoverage of the GATK were used to obtain various mapping statistics for unfiltered and filtered BAM files.

Mean effective sequencing depth, estimated from filtered BAM files, varied among individuals ranging from 4.8–12.2x [42] to 13.7–31.1x (this study) [43], with an average depth of 18.4x over all individuals (Tables S4). For the previously sequenced genomes [42, 43], estimated sequence depths were 25–40% lower as the values reported in the two source studies. This difference is explained by the way sequence depth was calculated. Here, we estimated sequence depth on the filtered BAM files where duplicated reads, bad read mates, reads with mapping quality zero, and reads which mapped ambiguously had already been removed. Thus, our sequence coverage estimates correspond to the effective read-depths which are available for SNP discovery and genotyping.

SNP and genotype calling

We produced high quality genotypes for all individuals for each position in the genome, applying the same filtering criteria for SNP and non-polymorphic positions. We identified SNPs and called genotypes in a three-step approach. First, we identified a set of candidate (raw) SNPs among all study individuals. Second, we performed variant quality score recalibration (VQSR) on the candidate SNPs to identify high-confidence SNPs. Third, we called genotypes of all study individuals at these high-confidence SNP positions.

Step 1: We used the *HaplotypeCaller* of the GATK in genomic Variant Call Format (gVCF) mode to obtain for each individual in the dataset genotype likelihoods at any site in the reference genome. *HaplotypeCaller* performs local realignment of reads around potential variant sites and is therefore expected to considerably improve SNP calling in difficult-to-align regions of the genome. We then genotyped the resulting gVCF files together on a per-island level, as well as combined for all individuals, using the *Genotype GVCFs* tool of the GATK to obtain three VCF files with candidate SNPs for *P. abelii*, *P. pygmaeus*, and over all *Pongo* samples.

Step 2: Of the produced set of candidate SNPs, we identified high-confidence SNPs using the VQSR procedure implemented in the GATK. The principle of the method is to develop an estimate of the relationship between various SNP call annotations (*e.g.* total depth, mapping quality, strand bias, etc.) and the probability that a SNP is a true genetic variant. The model is determined adaptively based on a set of 'true SNPs' (*i.e.*, known variants) provided as input. Our 'true SNPs' set contained 5,600 high-confidence SNPs, which were independently identified by three different variant callers in a previous reduced-representation sequencing project [62]. We ran the *Variant Recalibrator* of the GATK separately for each of the three raw SNP VCFs to produce recalibration files based on the 'true SNPs' and a VQSR training set of SNPs. The VQSR training sets were derived separately for each of the three raw SNP VCF files and contained the top 20% SNPs with highest variant quality score after having applied hard quality filtering as described for the VCF files in the BQSR procedure.

We used the produced VQSR recalibration files to filter the three candidate SNP VCFs with the Apply Recalibration walker of the GATK setting the '--truth_sensitivity_filter_level' to 99.8%. Finally, we combined all SNPs of the three VCF files passing this filter using the *Combine Variants* tool of the GATK, hence generating a master list of high-confidence SNP sites in the genus *Pongo*.

Step 3: We called the genotype of each study individual at the identified high-confidence SNP sites. We performed genotyping on the recalibrated BAM files in multi-sample mode for Bornean and Sumatran orangutans separately, producing one SNP VCF file per island.

Finally, we only retained positions with high genome mappability, *i.e.*, genomic positions within a uniquely mappable 100-mers (up to 4 mismatches allowed), as identified with the GEM-mappability module from the GEM library build [67]. This mappability mask excludes genomic regions in the orangutan reference genome that are duplicated and therefore tend to produce ambiguous mappings, which can lead to unreliable genotype calling. Furthermore, we aimed to reduce spurious male heterozygous genotype calls on the X chromosome due to *UnifiedGenotyper* assuming diploidy of the entire genome. We determined the male-to-female ratios (M/F) of mean observed heterozygosity (H_o) and sequence coverage in non-overlapping 20-kb windows along the X chromosome across both islands. We obtained a list of X-chromosomal windows where M/F of H_o was above the 85%-quantile or M/F coverage was above the 95%-quantile, resulting in 1255 20-kb windows requiring exclusion. We then repeated step 3 of the genotype calling pipeline on the X chromosome for the male samples setting the argument '-ploidy' of *UnifiedGenotyper* to 1 to specify the correct hemizygous state of the X chromosome in males. We subsequently masked all X-chromosomal positions within the spurious 20-kb windows in both male and female samples.

In total, we discovered 30,640,634 SNPs among all 37 individuals, which represent the most comprehensive catalogue of genetic diversity across the genus *Pongo* to date.

QUANTIFICATION AND STATISTICAL ANALYSIS

Recombination map estimation

We generated recombination maps for Bornean and Sumatran orangutans using the LDhat v2.2a software [68], following Auton et al. [69]. We used a high-quality subset of genotype data from the original SNP-calling dataset for the recombination map estimation for each island separately. Only biallelic, non-missing and polymorphic SNPs were used. Filtered genotype data were split into windows of 5,000 SNPs with an overlap of 100 SNPs at each side.

We ran the program *Interval* of the LDhat package for 60 million iterations, using a block penalty of 5, with the first 20 million iterations discarded as a burn-in. A sample was taken from the MCMC chain every 40,000 iterations, and a point estimate of the recombination rate between each SNP was obtained as the mean across samples. We joined the rate estimates for each window at the midpoint of the overlapping regions and estimated *theta per site* for each window using the finite-site version of the Watterson's estimate, as described in Auton & McVean [68].

We tested the robustness of the method with regards to the observed genome-wide variation of *theta* by contrasting recombination rate estimates using window-specific and chromosomal-average *thetas*. *Thetas* twice as large that the genome average produced very similar $4N_e r$ (*rho*) estimates. Because of this, a single genome-wide average of *theta per site* was used for all the windows (Sumatra: $\boldsymbol{\theta}_w$ = 0.001917, Borneo: $\boldsymbol{\theta}_w$ = 0.001309). We then applied additional filters following Auton et al. [69]. SNP intervals larger than 50 kb, or *rho* estimates larger than 100, were set to zero and the 100 surrounding SNP intervals (-/+ 50 intervals) were set to zero recombination rate. A total of 1,000 SNP intervals were found to have *rho* > 100 for *P. abelii*, and 703 for *P. pygmaeus*. In addition, 32 gaps (> 50 kb) were identified for *P. abelii*, and 47 gaps for *P. pygmaeus*. After applying the +/- 50 interval criteria, a total of 7,424 SNP intervals were zeroed for *P. abelii*, and 15,694 for *P. pygmaeus*.

Haplotype phasing

We phased the genotype data from Bornean and Sumatran orangutans using a read aware statistical phasing approach implemented in SHAPEIT v2.0 [70, 71]. This allowed us to obtain good phasing accuracy despite our relatively low sample sizes by using phasing information contained in the paired-end sequencing reads to support the statistical phasing procedure. We used a high-quality subset of genotype data from the original SNP-calling dataset containing only biallelic and polymorphic SNPs. We first ran the program extractPIRs to extract phase informative reads (PIR) from the filtered BAM files. In a second step, we ran SHAPEIT in read aware phasing mode using the following parameters: 200 conditional states, 10 burnin interations, 10 pruning interations, 50 main iterations, and a window size of 0.5 Mb. Additionally, we provided two species-specific recombination maps (estimated with LDhat) and the PIR files obtained in the first step to the program.

SHAPEIT uses a recombination map expressed in cM/Mb, therefore it was necessary to convert the LDhat-based *rho* estimates to cM/Mb units ($rho=4N_e r$). Accordingly, we estimated island-specific effective population sizes using the Watterson's estimator of *theta* (Sumatra: $N_e[\theta_W]=41{,}000$, Borneo: $N_e[\theta_W]=27{,}000$) and applied these to the recombination map conversion. The most likely pair of haplotypes for each individual were retrieved from the haplotype graphs, and recoded into VCF file format.

Individual heterozygosity and inbreeding

We determined the extent of inbreeding for each individual by a genome-wide heterozygosity scan in sliding windows of 1 Mb, using a step size of 200 kb. We detected an excess of windows with very low heterozygosity in the density plots, pointing to some extent of recent inbreeding. To estimate the cutoff values of heterozygosity for the calculation of inbreeding coefficients, we calculated heterozygosity thresholds for each island according to the 5th-percentile of the genome-wide distribution of heterozygosities (Borneo: 1.0×10^{-4} heterozygote sites per bp; Sumatra: 1.3×10^{-4}). Neighboring regions with heterozygosities below the cutoff value were merged to determine the extent of runs of homozygosity (ROH). Based on the number and size of ROHs, we estimated the percentage of the genome that is autozygous, which is a good measure of inbreeding [72]. We choose 1 Mb as window size for the calculation of heterozygosities based on previous studies identifying regions smaller than 0.5 Mb as the result of background relatedness, and tracts larger than 1.6 Mb as evidence of recent parental relatedness [73].

Sex-specific genomic data: mitogenomes and Y chromosomes

We produced complete mitochondrial genome (mitogenome) sequences for all study individuals. We first created a consensus reference sequence from 13 Sanger-sequenced mitogenomes representing almost all major genetic clusters of extant orangutans using BioEdit v7.2.0. [74]. The Sanger-sequenced mitogenomes were generated via 19 PCRs with product sizes of 1.0–1.2 kb and an overlap of 100–300 bp following described methods [75]. PCR conditions for all amplifications were identical and comprised a pre-denaturation step at 94°C for 2 minutes, followed by 40 cycles each with denaturation at 94°C for 1 minute, annealing at 52°C for 1 minute, and extension at 72°C for 1.5 minutes. At the end, we added a final extension step at 72°C for 5 minutes. PCR products were checked on 1% agarose gels, excised from the gel and after purification with the Qiagen Gel Extraction Kit, sequenced on an ABI 3130xL sequencer using the BigDye Terminator Cycle Sequencing kit (Applied Biosystems) in both directions using the amplification primers.

We individually mapped Illumina whole-genome sequencing reads of all 37 study individuals (Table S4) to the consensus mitochondrial reference sequence using NovoAlign v3.02. (NovoCraft), which can accurately handle reference sequences with ambiguous bases. This procedure prevented biased

951 short read mapping due to common population-specific mutations. For each individual, we generated
952 a FASTA sequence for the mitogenome with the *mpileup* pipeline of SAMtools. We only considered
953 bases with both mapping and base Phred quality scores ≥ 30 and required all positions to be covered
954 between 100 and 2000 times. Finally, we visually checked the sequence alignment of all individuals
955 in BioEdit and manually removed indels and poorly aligned positions and excluded the D-loop to
956 account for sequencing and alignment errors in those regions which might inflate estimates of mtDNA
957 diversity. In total, we identified 1,512 SNPs among all 50 individuals.

958 We thoroughly investigated the literature for the potential occurrence of nuclear insertions of mtDNA
959 (numts) in the genus *Pongo*, given that this has been a concern in closely related gorillas (*Gorilla*
960 spp.) [76]. There was no indication of numts in the genus *Pongo*, which is in line with our own
961 previous observations [23, 44, 45]. Numts also seem unlikely given our high minimal sequence depth
962 threshold.

963 We developed a comprehensive bioinformatics strategy to extract sequences from the male-specific
964 region of the Y chromosome (MSY) from whole-genome sequencing data. We expect the principle of
965 our bioinformatics strategy to be applicable to mammalian species in general if the taxon under
966 investigation is in phylogenetic proximity to one for which a Y-chromosomal reference sequence is
967 present or will be made available. Like for most mammals, there is currently no reference Y
968 chromosome for orangutans. Therefore, we had to rely on a reference assembly of a related species
969 (*i.e.*, humans) for sequence read mapping. Despite the ~18 million years divergence between humans
970 (*Homo* spp.) and orangutans [43, 77], we obtained a high number of MSY sequences. The impact of
971 varying Y chromosome structure among species [78, 79] on sequence read mappability might have
972 been reduced because we exclusively targeted X-degenerate regions. Hughes et al. [80] showed for
973 human and chimpanzees that although less than 50% of ampliconic sequences have a homologous
974 counterpart in the other species, over 90% of the X-degenerate sequences hold such a counterpart.

975 We applied several filters to ensure male-specificity and single-copy status of the generated MSY
976 sequences. (i) We simultaneously mapped sequencing reads to the whole orangutan reference genome
977 *PonAbe2* [42] and not just the human reference Y chromosome, reducing spurious mapping of
978 autosomal reads to the Y chromosome and allowing subsequent identification of reads that also
979 aligned to the X or autosomal chromosomes. (ii) We exclusively accepted reads that mapped in a
980 proper pair, *i.e.*, where both read mates mapped to the Y chromosome, which considerably increased
981 confidence in Y-specific mapping. (iii) We also mapped whole-genome sequencing reads of 23
982 orangutan females to the human Y reference chromosome and excluded all reference positions where
983 female reads had mapped from the male Y sequence data. (iv) To exclude potential repetitive regions,
984 we filtered non-uniquely mapped reads as well as positions with sequence coverage greater than two
985 times the median coverage for each individual, as extensive coverage can be indicative for repetitive
986 regions which might appear as collapsed regions on the Y reference chromosome. (v) To ensure that

we only targeted unique, single-copy MSY regions, we exclusively retained reads mapping to four well-established X-degenerate regions of the MSY in humans [81].

Our bioinformatics strategy consisted of the following detailed steps. First, we created a new reference sequence (*PonAbe2_humanY*) by manually adding the human reference Y chromosome (*GRCh37*) to the orangutan reference genome *PonAbe2* [42]. We then used BWA-MEM v0.7.5. [64] to map Illumina whole-genome short reads from 36 orangutans (13 males and 23 females) to this new reference sequence. We mapped reads for each individual separately in paired-end mode and with default settings. To reduce output file size, we removed unmapped reads on the fly using SAMtools v0.1.19 [82]. Picard v1.101 was used to add read groups and sort the BAM files. We then extracted all reads which mapped to the Y chromosome using SAMtools and marked read duplicates with Picard.

We used the GATK [65, 66] to perform local realignment around indels and filtered out duplicated reads, bad read mates, reads with mapping quality zero and reads which mapped ambiguously. We called genotypes at all sequenced sites with the *Unified Genotyper* of the GATK, applying the output mode 'EMIT_ALL_CONFIDENT_SITES'. We called genotypes in multi-sample mode (females and males separately, sample-ploidy was set to 1), producing one genomic VCF file for each sex. We only accepted bases/reads for genotype calling if they had Phred quality scores ≥ 30.

From the VCF file of the females, we generated a 'nonspec' list with the coordinates of all sites with coverage in more than one female (minimal sequence depth 2x), as these sites most likely were located in pseudoautosomal or ampliconic regions, *i.e.*, share similarity with the X or autosomal chromosomes. To ensure Y-specificity, we removed all sites of the 'nonspec' list from the VCF file of the males with VCFtools v0.1.12b. [83].

Finally, we used GATK to extract sequences of four well-established X-degenerate regions of the MSY in humans (14,170,438–15,795,786; 16,470,614–17,686,473; 18,837,846–19,267,356; 21,332,221–21,916,158 on the human reference Y chromosome assembly GRCh37/hg19)[81]. To be conservative, we chose regions which were longer than 1 Mb in humans and disregarded the first and last 300 kb of each region to account for potential uncertainties regarding region boundaries, leaving us with 3,854,654 bp in total. We exclusively retained genotype calls that were covered by a minimum of two reads and had a maximum of twice the individual mean coverage, resulting in 2,825,271 bp of MSY sequences among the 13 orangutan males. As expected, individual mean MSY sequence depth was about half (average: 54.4%) of that recorded for the autosomes, and ranged from 2.79–16.62x. For analyses, we only kept sites without missing data, *i.e.*, with a genotype in all study males. Because genomes of some individuals had been sequenced to only low coverage (~5–7x) [42], this left us with 673,165 bp of MSY sequences. We identified 1,317 SNPs among the 13 males, corresponding to a SNP density of 1 SNP every 511 bp.

We constructed phylogenetic trees and estimated divergence dates for mitogenome and MSY sequences using the Bayesian Markov chain Monte Carlo (MCMC) method implemented in BEAST v1.8.0. [50]. To determine the most suitable nucleotide substitution model, we conducted model selection with jModelTest v2.1.4. [52]. Based on the Akaike information criterion (AIC) and corrected AIC, we selected the GTR+I substitution model [84] for mitogenomes and the TVM+I+G model [85] for MSY sequences.

The mitogenome tree was rooted with a human and a central chimpanzee sequence from GenBank (accession numbers: GQ983109.1 and HN068590.1), the MSY tree with the human reference sequence *hg19*. We estimated divergence dates under a relaxed molecular clock model with uncorrelated lognormally distributed branch-specific substitution rates [86]. The prior distribution of node ages was generated under a birth-death speciation process [87]. We used fossil based divergence estimates to calibrate the molecular clock by defining a normal prior distribution for certain node ages. For mitogenomes, we applied two calibration points, *i.e.*, the *Pan-Homo* divergence with a mean age of 6.5 Ma and a standard deviation of 0.3 Ma [88, 89] and the Ponginae-Homininae divergence with a mean age of 18.3 Ma and a larger standard deviation of 3.0 Ma [77], which accounts for the uncertainty in the divergence date [90]. For MSY sequences, we used the Ponginae-Homininae divergence for calibration. We performed four independent BEAST runs for 30 million generations each for mitogenomes, with parameter sampling every 1,000 generations, and for 200 million generations each with parameter sampling every 2,000 generations for MSY sequences. We used Tracer v1.6 [91] to examine run convergence, aiming for an effective sample size of at least 1000 for all parameters. We discarded the first 20% of samples as burn-in and combined the remaining samples of each run with LogCombiner v1.8.0. [50]. Maximum clade credibility trees were drawn with TreeAnnotator v1.8.0. [50] and trees visualized in FigTree v1.4.0. [92] and MEGA v6.06. [93].

Autosomal genetic diversity and population structure

For all subsequent population genetic analyses, we assumed an autosomal mutation rate (μ) of 1.5 x 10^{-8} per base pair per generation, based on estimates obtained for the present-day mutation rates in humans and chimpanzees, derived primarily from de novo sequencing comparisons of parent-offspring trios but also other evidence [94-97]. There is good reason to believe that the mutation rate in orangutans is similar to that in other great apes, given the very similar branch lengths from outgroups such as gibbon and macaque to each species [98]. We assumed a generation time of 25 years [99].

We identified patterns of population structure in the autosomal genome by principal component analysis (PCA) of biallelic SNPs using the function 'prcomp' in R v3.2.2 [100]. Three separate analyses were performed: one within each island and one including all study individuals. For each sample set, we excluded all genotypes from the SNP VCF files that were covered by less than five

reads and only retained SNPs with a genotype call in all individuals after this filter. Furthermore, we removed SNPs with more than two alleles and monomorphic SNPs in the particular sample set. This restrictive filtering left us with 3,006,895 SNPs for the analysis of all study individuals, 5,838,796 SNPs for PCA within Bornean orangutans and 4,808,077 SNPs for PCA within Sumatran orangutans.

We inferred individual ancestries of orangutans using ADMIXTURE v1.23 [101]. We randomly sampled one million sites from the original VCF files and filtered this subset by excluding sites with missing genotypes or with a minor allele frequency less than 0.05. We further reduced the number of sites to 272,907 by applying a linkage disequilibrium (LD) pruning filter using PLINK v1.90b3q (–indep-pairwise 50 5 0.5) [102]. ADMIXTURE was run 20 times at all K values between 1 and 10. Among those runs with a difference to the lowest observed cross validation (CV) error of less than 0.1 units, we reported the replicate with the highest biological meaning, *i.e.*, runs that resolved substructure among different sampling areas rather then identifying clusters within sampling areas.

For subsequent analyses, we defined seven distinct populations based on the results of the PCA and ADMIXTURE analyses: three on Sumatra (Northeast Alas comprising North Aceh and Langkat regions, West Alas, and Batang Toru) and four on Borneo (East Kalimantan, Sarawak, Kinabatangan comprising North and South Kinabatangan, and Central/West Kalimantan comprising Central and West Kalimantan). Even though individuals from North and South Kinabatangan could be clearly distinguished in the PCA and ADMIXTURE analysis, we decided to pool the two Kinabatangan populations due to their low samples sizes (n = 2). This can be justified as data from the mitochondrial genome showed that they started to diverge only recently (~40 ka).

Ancestral gene flow between orangutan populations

We used D-statistics to assess gene flow between orangutan species, testing all three possible phylogenetic relationships among *P. abelii*, *P. tapanuliensis*, and *P. pygmaeus*. We extracted genotype data from the two individuals per population with the highest sequencing coverage and included two human genome sequences as outgroup (SRA sample accession: ERS007255 and ERS007266). We calculated D-statistics for all combinations of populations involving the three species using the qpDstat program of the ADMIXTOOLS package v4.1 and assessed significance using the block jackknife procedure implemented in ADMIXTOOLS.

To explore temporal patterns of gene flow between orangutan populations, we applied the multiple sequential Markovian coalescent (MSMC2) model [103]. The rate of coalescence of between-population haplotype pairs was compared to the within-population coalescence rate of haplotype pairs from the same population to obtain the relative cross-coalescence rate (RCCR) through time. A RCCR close to 1 indicates extensive gene flow between populations, while a ratio close to 0 indicates complete genetic isolation.

We used the phased whole-genome data for the relative cross-coalescence rate analysis. To avoid coverage-related issues, we selected the individual with the highest sequencing coverage for each population. We further excluded sites with an individual sequencing coverage less than 5x, a mean mapping quality less than 20, or sites with low mappability based on the mappability mask.

We ran MSMC2 for all pairs of populations, using a single individual (i.e. two haplotypes) per population. For each population pair, we performed three individual MSMC2 runs, using the default time discretization parameters: within population 1 (two haplotypes; -I 0,1), within population 2 (two haplotypes; -I 2,3), and between populations (four haplotypes; -I 0,1,2,3 -P 0,0,1,1). We then used the combineCrossCoal.py Python script of the MSMC2 package to combine the outputs of the three runs into a combined output file.

As the sequencing coverage of the best Batang Toru individual was substantially lower compared to individuals from other populations (~17x vs. ~23–27x, Table S3), we also assessed whether different sequencing coverage was negatively affecting the relative cross-coalescence rate results. To achieve this, we repeated the analysis using individuals with similar coverage as the Batang Toru individual (~16–21x). The results were highly consistent with the output from the runs with the highest-coverage individuals, indicating that the relative cross-coalescent rate analysis was robust to differences in sequencing coverage in our data set.

Approximate Bayesian Computation (ABC)

To gain insights into the colonization history of the Sundaland region by orangutans and obtain parameter estimates of key aspects of their demographic history, we applied a model-based ABC framework [26]. For this, we sampled a total of 3,000 independent sequence loci of 2 kb each, following the recommendations in Robinson et al. [104]. Loci were sampled randomly from non-coding regions of the genome, with a minimum distance of 50 kb between loci to minimize the effects of linkage. Since the coalescent simulations underlying ABC inference assume neutrality, we excluded loci located within 10 kb of any exonic region defined in the *Pongo abelii* Ensembl gene annotation release 78, as well as loci on the X chromosome and the mitochondrial genome, which would exhibit reduced N_e as compared to the autosomal regions.

For all ABC-based modelling, we defined three metapopulations for the calculation of summary statistics: Sumatran populations north of Lake Toba (NT), the Sumatran population of Batang Toru south of Lake Toba (ST), as well as all Bornean populations (BO). For each metapopulation as well as over all metapopulations combined, we calculated the first four moments over all loci for the following summary statistics: nucleotide diversity (π), Watterson's theta, and Tajima's D. Furthermore, for each of the three pairwise comparisons between metapopulations, we calculated the first four moments over loci of the number of segregating sites, proportions of shared and fixed polymorphism, average sequence divergence (d_{XY}), and Φ_{ST} [105]. To avoid potential problems with

unreliable phasing, we only used summary statistics that do not require phased sequence data. This resulted in a total of 108 summary statistics used in the ABC analyses. For each locus, we extracted genotype data of a total of 22 individuals (5 Northeast Alas, 5 West Alas, 2 Batang Toru, 4 Central/West Kalimantan, 2 East Kalimantan, 2 Sarawak, 2 Kinabatangan) by selecting the individuals with the highest sequence coverage for a given locus. Additionally, we recorded the positions of missing data for each locus and individual and coded genotypes as 'missing' in the simulated data if mutations fell within the range of missing data in the observed data.

In a first step, we used a model testing framework to infer the most likely sequence of population splits in the colonization history of orangutans. For this, we designed four models representing potential colonization patterns into Sundaland (Figure 3A). We assumed a simplified population structure with three distinct, random mating units composed of NT, ST, and BO metapopulations as described above. We simulated $4x10^6$ data sets for each model using the coalescent simulator ms [106]. Since we obtained a large number of summary statistics, we used a partial least squares discriminant analysis (PLS-DA) to extract the orthogonal components of the summary statistics that are most informative to discriminate between the four competing models using the 'plsda' function of the R package 'mixOmics' v5.2.0 [107] in R version 3.2.2 [100]. For model testing, we used the R package 'abc' v2.1 [108] to perform a multinomial logistic regression on the PLS transformed simulated and observed summary statistics, using a tolerance level of 0.05% (8,000 simulations closest to the observed data). To find the optimal number of PLS components for model selection, we performed cross-validations with 200 randomly chosen sets of summary statistics for each model and assessed model misspecification rates when using 10, 12, 15, 18, and 20 components.

We found that using the first 18 PLS components resulted in the lowest model misspecification rate. However, our model testing approach lacked power to reliably differentiate between pairs of models with the same underlying species tree (i.e. model 1a vs. model 1b and model 2a vs. model 2b in Figure 3A), as evidenced by a high model misspecification rate of 47.63% across all four models. In order to increase discrimination power with a new set of optimized PLS components, we therefore repeated the PLS-DA and multinomial logistic regression with the two best-fitting models (model 1a vs. model 1b). This resulted in a substantially lower model misspecification rate (36.00%). Moreover, no model misassignment occurred with a posterior probability equal or higher than the observed value (0.976), indicating a high confidence in the selected model (model 1a).

After establishing the order of population split events, we were interested in parameter estimates of different aspects of the orangutan demographic history. For this, we applied a more complex model that included additional population structure in NT and BO, as well as recent population size changes (Figure 3B). The design of this model was informed by (i) PCA and ADMIXTURE analyses (Figs. 2B and 2C), (ii) MSMC2 analyses (Figure 3C), and (iii) previous demographic modeling using more limited sets of genetic makers [49]. For parameter estimation, we performed a total of $1x10^8$

simulations as described above. Model parameterization and parameter prior distributions are shown in Table S5. We used 100,000 random simulations to extract the orthogonal components of the summary statistics that maximize the covariance matrix between summary statistics and model parameters using the 'plsr' function of the R package 'pls' v2.5-0 [109]. We defined the optimal number of partial least squares (PLS) components based on the drop in the root mean squared error for each parameter with the inclusion of additional PLS components [110]. After transforming both the simulated and observed summary statistics with the loadings of the extracted PLS components, we performed ABC-GLM post-sampling regression [111] on the simulations with the smallest Euclidean distance to the observed summary statistics using ABCtoolbox v2.0 [112]. To find the optimal proportion of retained simulations, we assessed the root-mean-integrated-squared error of the parameter posterior distributions based on 1,000 pseudo-observed data sets (pods) randomly chosen from the simulated data. We found that varying the tolerance level had little impact on the accuracy of the posterior distributions and therefore used a tolerance level of 0.00002 (equaling 2,000 simulations) for parameter estimation.

To assess the goodness of fit of our demographic model, we calculated the marginal density and the probability of the observed data under the general linear model (GLM) used for the post-sampling regression with ABCtoolbox [111]. A low probability of the observed data under the GLM indicates that the observed data is unlikely to have been generated under the inferred GLM, implying a bad model fit. We obtained a p-value of 0.14, showing that our complex demographic model is well able to reproduce the observed data. Additionally, we visualized the coverage of summary statistics generated under the demographic model relative to the observed data by plotting the first 12 principal components of the simulated and observed data. For this, we randomly selected 100,000 simulations and extracted PCA components using the 'prcomp' function in R. The observed data fell well within the range of simulated summary statistics for all 12 components. Furthermore, we checked for biased posterior distributions by producing 1,000 pods with parameter values drawn from the prior distributions. For each pods, we determined the quantile of the estimated posterior distribution within which the true parameter values fell and used a Kolmogorov-Smirnov in R to test the resulting distribution of posterior quantiles for uniformity. Deviations from uniformity indicate biased posterior distributions [113] and the corresponding parameter estimates should be treated with caution. As expected from complex demographic models, multiple parameters showed significant deviations from uniformity after sequential Bonferroni correction [114]. However, in most of these distributions, data points were overrepresented in the center of the histogram, which indicates that posterior distributions were estimated too conservatively.

G-PhoCS analysis

We used the full-likelihood approach implemented in G-PhoCS v1.2.3 [115] to compare different models of population splitting with gene flow and to estimate parameters of the best-fitting model. Due to computational constraints, we limited our data set to eight individuals with good geographic coverage of the extant orangutan distribution (1 Northeast Alas, 1 West Alas, 2 Batang Toru, 2 Central/West Kalimantan, 1 East Kalimantan, 1 Kinabatangan). We sampled 1-kb loci across the autosomal genome, ensuring a minimum distance of 50 kb among loci to minimize linkage. To reduce the impact of natural selection, we excluded loci located within 1 kb of any exonic region defined in the *Pongo abelii* Ensembl gene annotation release 78. We coded sites as missing based on the following filter criteria: low mappability, mean mapping quality less than 20, and individual coverage less than 5x. Sites without at least one valid genotype per species were excluded completely. We only retained loci with at least 700 bp of sites with data, resulting in a total of 23,380 loci for which we extracted genotype information for the eight selected individuals.

We compared models with the three different possible underlying population trees in a three taxon setting (Borneo, Sumatra north of Lake Toba, and Batang Toru). We performed 16 independent G-PhoCS runs for each model, running the MCMC algorithm for 300,000 iterations, discarding the first 100,000 iterations as burn-in and sampling every 11^{th} iteration thereafter. The first 10,000 iterations were used to automatically adjust the MCMC finetune parameters, aiming for an acceptance rate of the MCMC algorithm of 30–40%. We merged the resulting output files of independent runs and analysed them with Tracer v1.6 [91] to ensure convergence among runs. We then used the model comparison based on the Akaike information criterion through MCMC (AICM) [116, 117] implemented in Tracer to assess the relative fit of the three competing models.

In agreement with the ABC analyses, the model positing the deepest split between Sumatra north of Lake Toba and Batang Toru, followed by a split between south of Lake Toba and Borneo, showed a much better fit to the data compared to the two other splitting patterns. Independent replicates of the same model produced highly consistent posterior distributions, indicating convergence of the MCMC algorithm. All parameters of the best-fitting model were estimated with high precision, as shown by the small 95%-highest posterior density ranges (Table S5). Compared to the estimates from the ABC analysis, G-PhoCS resulted in more recent divergence time estimates for both the NT/(BO,ST) and BO/ST splits. This discrepancy might be caused by hypermutable CpG sites, which likely violate certain assumptions of the G-PhoCS model [115]. We could not exclude CpG sites in our analysis due to the absence of a suitable outgroup for calibration. Instead, we had to rely on a fixed genome-wide mutation rate, which includes hypervariable CpG sites. An alternative explanation could be a likely bias in the G-PhoCS results due to the restriction to a highly simplified demographic model as compared to our ABC analyses; G-PhoCS assumes constant effective population sizes and migration

rates in between population splits. However, this assumption is most likely violated in orangutans, as shown by the results of our ABC analysis (Figure 3B, Table S5).

Cranial, dental, and mandibular morphology

We evaluated five qualitative and 44 quantitative cranial, dental, and mandibular variables (Tables S1 and S2). We chose variables that had previously been used to describe and differentiate orangutan cranio-mandibular shape [14, 53, 54, 118-123]. Due to extensive dental wear of the Batang Toru specimen, we limited our comparisons with the Padang cave material to the breadth of the upper and lower canines, in addition to the length, breadth, and area (*i.e.*, breadth x length) of the lower first molar, all of which displayed a limited amount of wear. All measurements were taken by a single individual (AnN) in order to reduce observer bias.

We used both univariate and multivariate statistics to evaluate the Batang Toru specimen in relation to our comparative sample. As Batang Toru is only represented by a single sample, we first compared it to the interquartile range (IQR, defined as the range between the first and the third quartile) and the lower and upper inner fence (±1.5*IQR) for each separate sample population, using traditional methods for evaluating outliers [124]. This allowed us to evaluate the Batang Toru specimen's distance and direction from the central tendency of our sample orangutan populations. We also conducted univariate exact permutation tests for each morphological variable by removing a single sample for either the *P. abelii*, *P. pygmaeus*, or *P. p. palaeosumatrensis* sample populations and then comparing the linear distance to the mean of the remaining samples. This was done for each sample until all samples had a calculated value. A linear distance between the *P. tapanuliensis* sample and the *P. abelii*, *P. pygmaeus*, and *P. p. palaeosumatrensis* mean values (i.e., the test statistics) was then calculated and compared to the sample distributions detailed above. P-values represent the number of samples from the sample distribution that exceed the test statistic, divided by the total number of comparisons. In some cases, specimens did not preserve the measurements utilized in this study (*e.g.*, broken bone elements and/or missing/heavily worn teeth), and so were excluded from comparisons. Sample sizes for univariate comparisons of extant orangutan cranio-mandibular morphology are detailed in Table S1, whereas the sample sizes for the univariate comparisons of extant and fossil teeth are detailed in Table S2.

We also conducted a principal component analysis (PCA) on 26 of our 39 cranio-mandibular variables, on a subset of our extant orangutan sample, including *P. abelii* (n=8), *P. pygmaeus* (n=19), and the newly described *P. tapanuliensis* specimen. The choice of 26 variables allowed us to maximize sample size and avoid violating the assumptions of PCA [125]. A scree plot (using the *princomp* function from the base *stats* package in R [126]) indicated that seven principal components were sufficient to be extracted, based on the Kaiser criterion of eigenvalues at ≥1 [127]. Using the

principal function from the *psych* R package [128], we ran a PCA on the correlation matrix of our 26 selected variables, extracting seven principal components with varimax rotation.

To highlight the multivariate uniqueness of *P. tapanuliensis*, we used the extracted PCs and calculated the Euclidean D^2 distance for each sample relative to the *P. abelii* and *P. pygmaeus* centroids. We grouped these distances into two distributions, referred to as the between species (*i.e.*, the distances of all *P. abelii* samples to the *P. pygmaeus* centroid plus all of the *P. pygmaeus* samples to the *P. abelii* centroid) and within species (i.e, the distances of all *P. abelii* samples to the *P. abelii* centroid plus all of the *P. pygmaeus* samples to the *P. pygmaeus* centroid) distributions. We then compared the Euclidean D^2 distances of *P. tapanuliensis* to the *P. abelii* and *P. pygmaeus* centroids (i.e., the test values), relative to the two aforementioned sample distributions. Exact permutation p-values for these results were calculated as the number of samples from the sample distribution that exceed the test statistic, divided by the total number of comparisons. All Euclidean D^2 distance were calculated in the base *stats* package in R [126].

Acoustic and behavioral analyses

We used both previously published [129-131] and newly collected data in our analyses of male long calls. The current study includes n=130 calls from n=45 adult males across 13 orangutan field sites. In addition to two individuals from Batang Toru, we sampled 14 individuals of *P. abelii* and 29 individuals of *P. pygmaeus*. Using our comparative sample, we evaluated 15 long call variables (Table S3). We chose variables and their definitions that had previously been described to differentiate orangutan male long calls [129, 130, 132].

We used both univariate and multivariate statistics to evaluate the Batang Toru specimen in relation to our comparative sample. As Batang Toru is only represented by two individuals, we compared the mean of these two sample points to the interquartile range (IQR) and the lower and upper inner fence (±1.5*IQR) for each separate sample population [124]. As above, univariate exact permutation tests were conducted for each long call variable by removing a single sample for either the *P. abelii* or *P. pygmaeus* sample populations and then comparing the linear distance to the mean of the remaining samples. This was done for each sample until all samples had a calculated value. A linear distance between the average of the two *P. tapanuliensis* samples and the *P. abelii* or *P. pygmaeus* mean values (i.e., the test statistics) was then calculated and compared to the sample distributions detailed above. P-values represent the number of samples from the sample distribution that exceed the test statistic, divided by the total number of comparisons. In some cases, not all acoustic variables were available for each individual. As such, sample sizes for univariate comparisons are detailed in Table S3.

Geological and ecological analyses

We evaluated five ecological variables, including the type and age of geological parent material, elevation, average temperature, and average rainfall, to highlight that the current ecological niche of *P. tapanuliensis* is divergent relative to that of *P. abelii* and *P. pygmaeus*. For Sumatran populations, type and age of geological parent material were digitized from the land unit and soil map series of Sumatra [133-140]. No comparable geospatial data is available for Borneo, so we used previously published materials to more broadly characterize areas populated by orangutans [141]. To maintain consistency, elevation, average temperature, and average annual rainfall were collected from the WorldClim v. 1.4 bioclimatic variables dataset [142]. Using the digitized land unit/soil maps, we calculated the percentage of Sumatran orangutan distribution [143] classified into four classes for each type (*e.g.*, igneous, metamorphic, sedimentary, and other rock [*i.e.*, land units with a mixture of rock types]) and age (*e.g.*, Pre-Cenozoic, Tertiary, Quaternary, and other [*i.e.*, land units with a mixture of ages]) of geological parent material. For the elevation and climatic variables, we created 1km x 1km sample point grids for each currently identified orangutan population in Borneo and Sumatra [143, 144], and sampled the three aforementioned WorldClim datasets.

DATA AND SOFTWARE AVAILABILITY

Raw sequence read data have been deposited into the European Nucleotide Archive (ENA; http://www.ebi.ac.uk/ena) under study accession number PRJEB19688. Mitochondrial and Y-chromosomal sequences are available from the Mendeley Data repository under ID code doi:10.17632/hv2r94yz5n.1.